The Future Role of HPC in Medical Product Decision Making

Workshop Report
September 2015 Summary

CGSR

Center for Global Security Research

LLNL-TR-690736

LAWRENCE LIVERMORE NATIONAL LABORATORY

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Summary Report on a Workshop on the Future Role of High Performance Computing in Medical Product Decision Making

Abstract

On September 10, 2015, Lawrence Livermore National Laboratory's Center for Global Security Research[1] and High Performance Computing Innovation Center[2] hosted a workshop in Silver Spring, Maryland, on the "Future Role of High Performance Computing (HPC) in Medical Product Decision Making." Stakeholders from academia, industry, national laboratories and government agencies attended and discussed the role of HPC in making regulatory decisions. This document summarizes those discussions. It presents the benefits and challenges of the use of data driven discovery and computationally intensive modeling and simulation that were identified by attendees, as well as potential courses of action that would support the acceptance and the adoption of HPC-backed regulatory science. The participants concluded that HPC can clearly enable regulatory science, but several challenges must be addressed:

- There is a critical need for personnel.
- There is a need for an enterprise level business model for big data health analytics.
- Confidence building measures must be proposed and initiated early in the process.
- Expand the role of uncertainty quantification in the biological sciences to ensure a rigorous and well-documented method of defining confidence levels.
- There is an urgent need for a community wide forum to share ideas, best practices and results.

Introduction

Twenty-five years ago, *Science* magazine declared the debut of the third branch of science—a means of performing "experiments" that would otherwise be prohibitively expensive or simply impossible to execute. In the subsequent quarter century, computer power has increased 100,000 fold and massive databases aided by the ubiquitous connectivity provided by the Internet have come into being. Now the "Fourth Paradigm" of science—data driven discovery—is upon us.

The objective of regulatory science is to instill trust and confidence in the products being regulated. Computer models and simulations (see the sidebar for an explanation of the difference between models and simulations) are intended to replace expensive and infeasible physical experiments. Historically, confidence in computer models and simulations was provided by a reliance on the scientific method—

[1] Lawrence Livermore National Laboratory - https://cgsr.llnl.gov/
[2] Lawrence Livermore National Laboratory - http://hpcinnovationcenter.llnl.gov/

models and simulations were backed up both by experimentation, confirmation by multiple research teams and peer review. By contrast, confidence building measures are exceedingly difficult and often notably absent when one discusses the third branch and the fourth paradigm, both enabled by HPC. Is it possible for modeling and simulation to significantly supplement or even replace clinical trials without negatively impacting the public's trust? Determining how much trust to invest in HPC "black boxes" is a complex technical and social problem. If the nation is to tap into the enormous potential benefits of HPC and advanced models and simulations, we must solve the intertwined issues of uncertainty, data integrity and trust.

The objective of this workshop was to inform the different communities of interest of both the benefits and pitfalls of using HPC for regulatory science. By initiating a conversation among all stakeholders as HPC capabilities mature, the technical capabilities and regulatory requirements might be developed in parallel. "One of the confounding issues in translating a novel discovery into clinical practice is that quite often the scientists working on… discovery have limited knowledge of the analytical, diagnostic,\ and regulatory requirements for a clinical assay."[3] This integrated approach would also provide guidance to computational scientists so that the products they help develop will be compliance ready. With advanced planning, one might ensure that the regulatory process will capitalize on HPC benefits on a timely basis.

The workshop was structured in two parts (workshop agenda, appendix A). The morning session commenced with two keynote presentations (appendices B and C), which provided a background for participants on current computational efforts at the Food and Drug Administration (FDA) and the state of the art in HPC. These were followed by three panel sessions (appendix D) which provided further background in the areas of big data, simulations and applications of HPC to animal models.

The second half of the workshop was comprised of two breakout sessions that examined big data analytics and mechanistic models. The day concluded with a short session to summarize the day's discussions.

Background on Medical Product Modeling & Simulation – Capabilities and Application (morning session)

Modern super computers are five orders of magnitude more powerful than desktops. To put this in perspective, compare the historical game PAC-MAN (run on a then state of the art desktop computer) to current three-dimensional, full physics-based, real-time animated multi-person games that run on today's desktop (about 100,000 times faster than early PCs). Now extrapolate by another factor of 100,000 from the modern desktop to a supercomputer. The simulation of a biological pathway which runs on a workstation can be transformed into a full, mechanistic simulation of an organ when implemented on a supercomputer.

[3] Fuzery et. al.; "Translational of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges," Clinical Proteomics 2013, 10:13.

Experience tells us that today's world class computing capabilities will be widely available in industrial settings in 5–10 years, a time frame comparable to a drug or medical device development cycle. Having this level of capability means it is or will soon be possible to:

- Simulate an entire organ rather than using mechanistic models of biological pathways. A current heart simulation[4] uses 700 million volume elements permitting one to study drug safety and to understand mechanisms of cardiotoxicity well beyond single-cell effects.
- Replay computational experiments thus permitting rapid exploration of many parameters, including human variations.
- Investigate cooperative effects of the use of multiple drugs or treatments.
- Investigate the utility of different medical products. For example:
    - Optimization of Cardiac Resynchronization Therapy
    - Tissue-lead interface models and simulations
    - New ablation techniques and instruments including electromagnetic effects
    - New treatment approaches for atrial fibrillation

For the purposes of this workshop summary we utilize the following definitions:
Model - A collection of calculations and algorithms utilized for analyzing test data, making statistically based inferences, and machine learning.
Simulation - A computationally based reproduction of a process or system based upon physical principles.
Mechanistic model – A term used interchangeably with simulation.

Consequently, the next generation of medical products may be developed on corporate-owned high-performance computers with capabilities that match or even exceed today's highest performing machines.

The FDA realizes the potential benefits of utilizing HPC to meet their regulatory responsibilities, and is studying how models and simulations may beneficially enable the regulation of medical products. Their principle computational initiatives are designed to exploit the opportunities presented by the availability of multiple large databases—big data. The opportunity for large advances clearly exists because:

- Large and diverse databases are widely and easily available to the public and corporate enterprise.
- Software tools are often available on the web.
- Open databases on labelling, adverse medical product effects, genomic sequences etc., will be growing and increasingly available.

Applying advanced analytics to large databases could potentially inform many of FDA's key determinations, such as the safety, effectiveness, manufacturing quality, or legality of the product, as well as where it is manufactured (domestically or internationally). Current and future databases might

[4] LLNL heart simulation needs a reference

contain not only clinical data, but also physical data (e.g. pictures of the product or even views of a manufacturing site obtained by drone facilitated inspections) and trade data (e.g. sales records, supply chains).

Public availability of the medical product data increases transparency (and consequently trust) between and among all stakeholders—regulators, industry and the public. The FDA has taken the initiative by making much of their data accessible in several open databases. Current FDA initiatives[5] include:

1. China safety initiative which develops dashboards and models to predict unsafe or ineffective imports by using data from trade, Chinese FDA, Chinese media and FDA.
2. precisionFDA is an online, cloud-based portal that will allow scientists from industry, academia, government and other partners to come together to foster innovation and develop the science behind a method of "reading" DNA known as next-generation sequencing.
3. GenomeTrakr network and Chillax—GenomeTrakr consists of public health and university laboratories that collect and share genomic and geographic data from foodborne pathogens. The data, which are housed in public databases at the National Center for Biotechnology Information, can be accessed by researchers and public health officials for real-time comparison and analysis.
4. Sentinel aims to build and implement an active surveillance system that will eventually be used to monitor all FDA-regulated products.
5. Open FDA illustrates how HPC can increase transparency. "FDA's public databases and a number of open source analytics tools are hosted by Amazon web services. The concept is to index high-value public-access data that is formatted and documented in developer and consumer-friendly standards, and then make that data available via a public-access portal that enables developers to quickly and easily use it in applications."[6]

There is some precedent for utilizing computational models and simulations in regulatory decision making. For over two decades, the FDA has accepted certain types of non-clinical drug evaluations for regulatory purposes. Potential advantages of these non-clinical methods (including stem cell assessments) include:

- Better hazard identification, risk assessment and translation
- The ability to more easily investigate sub populations and natural variability
- The potential to mitigate the hERG effect[7] and reduce the number of inappropriately discontinued trials of promising drugs

Several issues and open questions requiring further exploration were identified during the morning sessions by the participants:

[5] See appendix B
[6] https://open.fda.gov
[7] Human ether-a-go-go-related gene

- Open data is being accessed principally by simple (non-compound) queries. There appeared to be few examples of advanced analytics that combine multiple databases to answer complex queries or employ machine learning techniques.
- The data exists in multiple formats and types (images, text, audio and instrumental). Conversion to compatible formats needs to be automated. This is critical if data in multiple formats is to be successfully combined, mined and fused.
- Hardware, software and data scientists are in short supply. The community of data analytics specialists with knowledge of medical products must be greatly enlarged if in silico testing is to become a basic tenet of regulatory science.
- Enhanced collaboration between scientists, clinicians and computer scientists is required.
- Regulators will need to identify what information is needed to validate the models and simulations. In addition, industry, the public and regulators must all be confident and comfortable with the "black box."
- An exemplar case would be beneficial. HPC efforts in cardio vascular pharmacology may serve as a test case e.g. for Pro-arrythmia Risk assessment (CiPA) through scientific evidence will have to be presented that demonstrates that HPC enabled regulatory approval is as effective as previous methods.
- We need to define endpoints (what constitutes a beneficial therapeutic result?)—traditional or other.
- For HPC-derived data to supplant (not merely support) current trials for clinical and non-clinical safety assessment. the level of reliability on HPC-enabled models and simulations must be determined. The commercial sector will have to be comfortable with whatever standards are eventually implemented.
- Models and simulations must be fit for purpose, though it remains unclear what fit for purpose requires.

Further discussion: Big Data Analytics and Mechanistic Models (afternoon session)

Modeling and simulation for medical product regulation naturally divides into two categories—Big Data Analytics and Mechanistic Models.

- Big data analytics and statistical models – drug effects extrapolated from near neighbor molecules and chemical compounds. These often rely upon biomarkers as indicators of therapeutic effectiveness. The models are correlation based, often with minimum or less than desirable understanding of causation.
- Mechanistic models – simulations of biological systems based upon physics, chemistry and biology. The biological mechanisms are simulated with detail and resolution dependent upon both computational resources and understanding of the underlying processes. Examples include simple metabolic or pathway mechanistic models, cellular mechanistic models of the beating

heart, predicting the binding of chemical entities to potential on–and–off target receptors and the transmission of efficiency of physical signals through tissues. These simulations are generally causality based, although the underlying biology is often incomplete and uncertain.

Mechanistic models illuminate the underlying processes, but incomplete biology and natural variation among individuals leads to uncertain outcomes. Big data analytics can point to favorable biomarkers correlated with other successful approaches, but with only a weak understanding of causation and therefor again uncertain outcomes.

As HPC capabilities become commonplace, both approaches will be challenged to demonstrate that computationally enabled decisions are backed by strong scientific evidence. The methods that might be utilized to provide confidence in both mechanistic models and big data analytics have some similarities (transparency, repeatability, the need for uncertainty analysis and the ability to deal with natural variability) and some differences (physical fidelity vs. statistical justification and the accuracy and curation of the underlying "big" public and proprietary databases).

Big Data Analytics

According to the National Science Foundation, the phrase big data "refers to large, diverse, complex, longitudinal, and/or distributed data sets generated from instruments, sensors, Internet transactions, email, video, click streams, and/or all other digital sources available today and in the future."[8] Big data in health care might include results of longitudinal studies, medical product information, electronic health care records, patient data, x rays and other instrumental data, social data and even mobile data. McKinsey Global Institute estimates that applying big data to better inform decision making in health care could be worth $100 billion per year.[9] The participants discussed the obstacles that must be overcome if the full potential for big data in support of medical product decision making is to be realized.

The points highlighted during the discussion of big data were as follows:

1. ***There is a need to formulate and promulgate standards for data (both reference data and "fluid" data—data generated by users of big data analytics), metadata, actionable knowledge and bioinformatics.***

   Some outstanding issues identified are as follows:

   a. Analytics engines may need their quality control standards if they are used to access the quality of big data and metadata.

[8] National Science Foundation program solicitation for Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA), June 2012; http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.pdf

[9] McKinasey and Company – http://www.mckinsey.com/insights/health_systems_and_services/how_big_data_can_revolutionize_pharmaceutical_r_and_d

b. A suitable authorized organization (either the government with full regulatory authority or a consortium in an advisory capacity) should set and maintain these standards.
c. This function may be part of an enterprise business model (see point 6).

Nomenclature, ontology, methods of normalization and data formats need to be standardized to help realize database interoperability. The quality of public database data varies, and there is a need for a curation, quality control and clear exposition of data provenance.

2. ***Ensuring privacy while maximizing the utility of the data was considered to be a major challenge.***

Some outstanding issues identified are as follows:

a. The demographic distribution (young vs. old, veterans vs. civilians, etc.) of those willing to share health data requires consideration when applying big data to regulatory actions.
b. The utilization of different data types (genetic, electronic health records, personal data and mobile data) might benefit from having different, data-specific consent forms.
c. Access to and ownership of such patient data by FDA, the National Institute of Health, pharma, hospitals, medical researchers and patients themselves must be determined.
d. The Health Insurance Portability and Accountability Act of 1996 (HIPAA) and the Common Rule need to be reexamined and modified so that the U.S. regulatory environment can be made consonant with the era of big data analytics.

Current European regulators are apparently moving towards open health consents (opt out) while the U.S. seems to be moving in the opposite direction (opt in).[10] Integrating research data with clinical data could be particularly challenging when the need for anonymization and de-identification is considered. The details will depend on the chosen infrastructure and architecture (see point 6).

3. ***There is a need to disseminate success stories.*** In order to advance discussion, gain the public's confidence, and prove the ability for models to provide evidence of safety and efficacy to stakeholders, examples of expedited product delivery and reduced development costs should be promulgated.

4. ***The lack of trained information scientists and professional software development processes (codes are often developed by students and never documented or maintained) was acknowledged as a major stumbling block.*** Instituting formal exchange programs whereby professionals could for brief periods move between intuitions—academia, government and industry—was viewed as an excellent means by which data scientists could gain experience in all

[10] It was pointed out that this seems to be a reversal of the two positons on internet privacy

facets of regulatory science and HPC, thereby appreciating the issues native to other organizations.

5. ***It was noted that different data business models are utilized within the USG.*** NIH follows an integrated model organized around big data centers with significant sharing across components while the FDA is vertically integrated. It was suggested that the FDA might benefit from having a horizontal bioinformatics capability.

6. ***Many of the issues can be consolidated within the context of a need for an enterprise level business model.*** Options included a federated database[11] versus a single centralized database. The federated data system reduces the complications that accompany moving and securing large databases, while centralized systems have a computational advantage if large HPC platforms are required. Federated systems also help ameliorate ownership options while centralized systems might exacerbate them if economies of scale lead to a small number of large commercial systems resulting in limited competition and increased costs to database clients. A public–private partnership in which the government seeds the effort with data and academia and industry provide community based analytics is another possible option. The FDA (medical devices) is moving in this direction. An associated issue is data curation—clearly federated and centralized systems—will need different procedures to document the provenance and assure the integrity and accuracy of data. Community-based algorithms will also need to be curated and tested. One method is to test algorithms against standardized "golden data" databases before testing them on real data. The scope of algorithmic validation (pipeline vs. algorithmic) may differ depending on the user base and needs. A business model would provide a structured framework to formulate and analyze these options.

Mechanistic models

A mechanistic model of a medical product is a simulation of physiological function based upon physical principles—kinetics, fluid transport, mechanics, electro-magnetics, chemical potentials, material properties, to name a few. Mechanistic models have had enormous success in simulating complex systems, everything from airplanes to nuclear weapons, and are beginning to be accepted by the FDA as providing evidence of the safety and efficiency of medical devices.[12] However, the totality of systems within the human body comprise perhaps the most complicated "system of systems." Many of these underlying processes and pathways are ill defined and unknown. Nevertheless, rapid progress is being made to understand these systems, and the potential impact of HPC on physiological simulations is enormous. The participants considered issues related to the use of mechanistic models to reduce approval costs and expedite regulatory processes while preserving and enhancing the safety of medical

[11] "A federated database system is a type of meta-database management system (DBMS), which transparently maps multiple autonomous database systems into a single federated database. The constituent databases are interconnected via a computer network and may be geographically decentralized." Wikipedia

[12] The MRI compliant pacemaker is a case in point

products. The undetermined future of HPC-based mechanistic models in medical product regulatory scienceis an outstanding issue. These simulations may also impact the use of animal models and clinical trials.

The points highlighted during the breakout sessions were as follows:

1. ***The potential benefits of HPC to enable physiological simulations are many and varied*.** The rapidity and ability to repeat trials and test hypothesizes across many and varied parameters, populations and sub-populations were often mentioned as a critical advantage of in-silico testing. Early screening for adverse drug effects by simulations could reduce the size and cost of clinical trials leading to accelerated development. This could be an advantage for identifying and qualifying biomarkers.

   Mechanistic models may also be particularly suited for extrapolating product effects from narrow trial populations to a more representative and complete demographic (however, see cautions below). Simulations could also help explore a drug's impact across the full natural variability of the human species, capturing the tails of the population distribution function, which is often undetected in even the largest and costliest clinical trials. These simulations might help ameliorate the known difficulty (and consequent risk) of moving from animal models to humans—human disease pathways differ significantly from those of animals. Finally, HPC-supported mechanistic models might enable scientists to extrapolate from results obtained on well characterized populations to those about which there is a paucity of data.

   The simulations are not only useful for broadening the scope of testing, but they could also be critical in the development of personalized, individually targeted medicine. One could imagine introducing the mechanistic models into a clinical setting—creating a customized virtual heart for every patient leading to patient specific strategies and treatments.

2. ***Along with the benefits, workshop participants identified a multitude of cautions that must be considered as the community attempts to incorporate massive mechanistic models into the regulatory process.***

   Some outstanding issues identified are as follows:
   a. Though healthy heart simulations have proved accurate, these mechanistic models may not represent unhealthy hearts well.
   b. In the existing simulations some pathologies are well known, but lesser-known conditions may not be described.
   c. Many simulations predict typical or average heart behaviors but it is not known how they will represent behavior variability within large populations.
   d. These simulations provide time-critical data but the complexity of these tools may hinder potential users unfamiliar with HPC.

Many of the benefits listed above can also be considered reasons to be cautious. Recently, a number of extensive healthy heart simulations have been developed, but these simulations may not accurately represent unhealthy hearts. Part of the difficulty is the very nature of biological systems and the number of unknowns. Current simulations can be over parameterized for an environment with so many unknowns—pharmacokinetic–pharmacodynamic (PK–PD) simulations may have 20–30 parameters. There are too many "knobs" without sufficient data to accurately set them. The simulations must deliver reproducible results in a reasonable amount of time. Given the technical complexity of HPC, users unfamiliar with HPC may experience difficulty operating these tools.

Finally, although it is essential to follow the scientific method carefully, it is not clear how that might be accomplished when one is trying to validate human physiological simulations. There is currently, and there is likely to remain for some time much less validating data is available for these simulations than is typically available for physics or engineering mechanistic models. More might be done with the data currently available if there was more cooperation within the community (see below). Even the verification (confirmation of proper implementation and execution of the specified algorithms) of a simulation running on a high-performance computer is a highly specialized and complex task.

3. ***There is a need for more coordination between clinicians, experimental biologists and mechanistic model developers.***

Some outstanding issues identified are as follows:

a. The most effective scale for mechanistic models and confirming experiments, local or global scale (i.e. simulating the transmission of an electromagnetic wave across the heart or searching for regional heart disruptions) is unclear.
b. Mechanistic models could be made to help simulate single systems or to understand cross-system interactions, though it is undetermined which would be most applicable.
c. Mechanistic model developers must make their needs for validating experiments and data clearly known to industrial scientists who may possess the data.

The scientific method is rooted in the concept of conducting experiments specifically designed to test hypothesizes. When theories concern very complicated systems with scores of confounding factors, this can be a difficult proposition. Typically, the complexity is tackled by simplifying early experiments and increasing complexity as confidence in the simulations and hypothesis is gained, but to date precious few experiments have been conducted with the goal of guiding and validating physiological simulations.[13] These considerations prompted a discussion of the appropriate scale for mechanistic models and confirming experiments. HPC

[13] For example might you confirm in silico simulations with stem cell derived myocytes?

simulations should be able to guide the design of validating experiments, and they themselves should be constructed with the need for validating experiments as a requirement.

Though a large amount of potentially applicable diseased animal model information is available, the available data is not yet being extensively utilized for simulation verification. The extensive body of research with rabbit models might be a first step to simulation validation and testing, although the experiments that have been conducted have not been designed with the express intent of validating simulations. Given the lack of data and the wide range of only marginally known parameters, it was suggested that there could be a large role for uncertainty quantification (UQ) in the verification and validation of medical product mechanistic models. A first step is to create a simulation framework that incorporates a UQ level of rigor, beginning with physics-based models. This framework will require input from industrial and academic researchers. UQ is likely to be critical in a regulatory framework relying on HPC-based simulations, however UQ applied to biological systems is in a nascent state.

More coordination and communication between the simulation developers and the experimental community would help address the lack of validation data. Communication between experimentalists, mechanistic model developers, and industrial scientists could make more data available to simulators.

Part of the problem may be that there is not yet a critical mass of scientists and organizations with the skill to support HPC-based physiological modeling. It might help if simulations were developed that could be utilized by non-HPC scientists. As code developers build their simulations, this should also be considered a requirement.

4. ***The appropriate use for HPC-enabled simulation at each stage of the "womb-to-tomb" process is unclear.***

Some outstanding issues identified are as follows:

a. When following steps associated with medical product regulatory approval, it is essential to determine at each step the minimum validation requirements, when HPC could be used to augment clinical data, and how these considerations might evolve over time as simulations advance and confidence in in silico testing increases.
b. Particular steps in the validation process may be best matches to in silico based decisions.

The regulatory process is carefully proscribed with many steps to be completed before (and even after) a medical product is approved. For initial attempts, a specific focus using HPC on a single step may help build confidence in the benefits of simulation. One suggestion was to use biomarker qualification as low hanging fruit that might both show the advantages and issues that need to be resolved. Another might be to demonstrate the ability to predict the behavior

of a simple physiological subsystem that can be expanded to more complex systems. An example would be the current efforts to simulate the propagation of electrical currents through the ventricles, eventually building a full human heart that provides measures of force and pressure that reproduce known drug effects.

5. ***Ultimately, the success or failure of using HPC to support medical product decisions is all about confidence—confidence of the regulators in the evidence submitted by the regulated, confidence of the scientists (modelers, testers and other health professionals) in the simulations and public confidence in the regulators.***

   Some outstanding issues identified are as follows:

   a. Confidence in using HPC to validate medical research must be gained without undermining confidence in previous procedures.
   b. The most capable person(s) to present HPC's benefits and reduced risks must be identified.

   The question becomes one of who should make the case for using HPC to support medical applications and how it would be best accomplished. Convincing the technical community may require more rigor than the public or even health professionals. What lessons might be learned from in silico testing in the aviation or nuclear weapons sectors? Maximum transparency and data sharing is certainly required, but those who are privy to this information—regulators, independent peer reviewers or competitors—must be identified. Parts of simulations must be left opaque to preserve IP.

   A paradigm shift will be required, either evolutionary or revolutionary. Many workshop attendees felt a gradual approach to confidence building was best; one should carefully choose the context of use of the first simulations and products. An early and continuous dialogue between all parties (researchers, product developers and clinicians) and the FDA is essential if first steps are not to become the last.

Summary:

HPC can support and enable regulatory science. The benefits are many and substantial, but in order for them to be realized in a timely manner, several challenges must be met:

1) There is a need for manpower to support these developments.

a. Data analytics personnel with knowledge of medical products and regulatory science are in short supply. An exchange program ("sabbaticals" or IPAs[14]) between the regulatory agencies, academia and industry would start the process.
b. Researchers, clinicians and computer scientists need to engage in more communication and collaboration. Annual workshops are necessary but not sufficient. A continuing (perhaps web-based) community of interest for HPC-enhanced regulatory science should be established.

2) There is a need for an enterprise-level business model for big data health analytics.
   a. The pros and cons of a federated or integrated system must be evaluated.
   b. Privacy and ownership issues must be addressed—HIPPA and the common rule need to be reexamined in a world of big data analytics.
   c. Data formats need to be standardized—who is in charge? Which organization—government, non-governmental, or even commercial—should take the lead?
3) Confidence building measures must be proposed and initiated early in the process.
   a. Start slowly—find some "low hanging fruit" and expand. One possibility is qualifying biomarkers or demonstrating the ability to predict behavior of a simple physiological subsystem that can be expanded to more complex systems.
   b. The early uses of HPC have to be at least as good (safety and efficacy evaluations) as current methods. Where in the "womb-to-tomb" regulatory process is the appropriate starting point for HPC-based mechanistic models?
   c. Success stories need to be widely disseminated in order to advance the discussion and gain confidence. Transparency and professional debate among scientists is essential.
   d. Follow the scientific method to establish an evidence-based protocol and maintain confidence in the results. This will require purpose-designed experiments for hypothesis testing and validation.
   e. Expand the role of uncertainty quantification in the biological sciences to ensure a rigorous and well documented method of defining confidence levels.

Acknowledgements:

This report summarizes the content of a workshop held on September 10, 2015. All the participants of workshop contributed, but the comments are not necessarily the specific opinions of any individual and should not be attributed to any individual. We would like to especially thank members of the LLNL team, Ken Turteltaub, Felice Lightstone, Michele Pearson, and Laura Schultz; our keynote speakers, Taha Kass-Hout and Fred Streitz; and our panel facilitators and leads, David Strauss, Vahan Simonyan, and Hugo Vargas for stimulating and leading the discussions.

[14] Intergovernmental Personnel Act

Paris Althouse and Don Prosnitz, editors

Appendices

A- Agenda
B- Keynote: Taha
C- Keynote: Streitz
D- Panel presentation: Vargas

Appendix A

Lawrence Livermore National Laboratory

Workshop: **What is the Future Role of HPC in Medical Product Decision-making?**

Silver Spring Civic Building, Silver Spring MD
September 10, 2015

Advisory Committee Members

Paris Althouse LLNL

Felice Lightstone LLNL

Don Prosnitz LLNL

Fred Streitz LLNL

Ken Turteltaub LLNL

Brian Fitzgerald FDA

Tina Morrison FDA

Antonio Paredes FDA

Vahan Simonyan FDA

Jack Reynolds Anabios

Richard Arthur GE Global Research

Russell Thomas EPA

Don Bers UCD

Time	Workshop topic	Presenter/Chair
8:30 - 9:00 am	Welcome/Purpose /Introductions	Don Prosnitz, LLNL
9:00 – 9:30	FDA Key Note on Regulatory Science	Taha Kass-Hout, FDA
9:30- 10:00	Key Note on Future of HPC in Drug Development and Health	Fred Streitz, LLNL
10:00- 10:30	Briefing/Panel on Topic 1 (**Simulation**)	Facilitator: David Strauss, FDA
10:30- 10:45	BREAK	NA
10:45- 11:15	Briefing/Panel on Topic 2 (**Big Data Analytics**)	Co-Facilitators: Vahan Simonyan, FDA Taha Kass-Hout, FDA
11:15- 11:45	Briefing on Topic 3 (**How will HPC impact the use of clinical trials?**)	Facilitator: Hugo Vargas, Amgen
11:45- 12:45	LUNCH (provided by LLNL)	ALL
12:45- 1:00	Structure and Goal of Breakouts	ALL
1:00- 3:00	Breakout discussions	Facilitators
3:00- 4:00	Wrap up/ Breakout Read Out for attendees and others	Breakout chairs
4:00- 4:15	Workshop Summary and Path Forward	

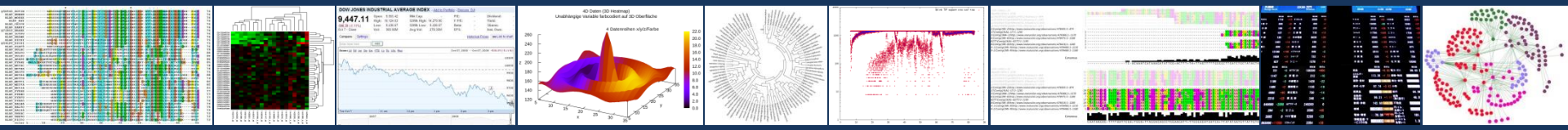
Appendix B

FDA: Future Role of HPC

Taha A. Kass-Hout, MD, MS
Chief Health Informatics Officer
Roselie A. Bright, ScD, MS, PMP
both in FDA's Office of Health Informatics

Keynote for Workshop: Future Role of High Performance Computing (HPC) in Medical Product Decision Making

9/10/2015



Outline

- HPC
- Current FDA activities that could be advanced with HPC
- State of these HPC activities
- What would it take to enhance current FDA activity with an HPC-enabled activity?

Important Disclaimer: this presentation has not been cleared as official FDA policy and may be incomplete. It is merely intended to spark ideas and discussion.

High Performance Computing (HPC)

- HPC involves
  - Hardware: more nodes, faster, bigger
  - Software: faster, shortcuts
    - Open source: mix and match modules, clever
  - Data scientists in short supply
- HPC enables
  - Speed
  - Big data (no sampling!)
  - Complex algorithms
  - “Citizen” analysts

HPC Resources on Internet

- Data: GenBank
- Application Programming Interfaces: openFDA
- Computation sites: http://app.raw.densitydesign.org/#%2F
- Downloadable free, open code
  - Pipelines: Python Luigi
  - Analytics: R
- Generally no expectation of privacy
  - PrecisionFDA will allow user to select level

Future of HPC Resources on Internet

- Patient-provided shared data (beyond electronic health records) will grow in amount and relevance.
- Open reference databases will be more relevant:
  - Registration and Listing
  - Labelling
  - Substances
  - Body components and functions
  - Health conditions
  - Whole Genome Sequences
  - ClinVar/Gen
  - Current knowledge
  - Untested hypotheses
- Linkages across databases will exponentially increase computing demand.
- To deal with text, signals, and images will need to either:
  - Automatically code it to standards, or
  - Automatically use its meaning without standardization

FDA activities that could be enhanced with HPC

Relative to medical products (MP), FDA mission is to protect and promote public health by advancing regulatory science

FDA activities revolve around questions:

1. **Who makes MP, where (US)?**
2. **Who makes MP, where (non-US)?**
3. **Is MP legal?**
4. **Is MP labelling legal?**
5. **Is MP made properly?**
6. **Is MP safe?**
7. **Is MP effective?**
8. **Is new moderate risk MP like a legal MP?**
9. **Is new risky MP safe and effective?**

FDA activities to address the questions

Questions

1. Who makes MP, where (US)?
2. Who makes MP, where (non-US)?
3. Is MP legal?
4. Is MP labelling legal?
5. Is MP made properly?
6. Is MP safe?
7. Is MP effective?
8. Is new moderate risk MP like a legal MP?
9. Is new risky MP safe and effective?

Activities (all pre-date computing)

- Manufacturers register, and list their MP, with FDA
- FDA inspects manufacturer's site
- FDA inspects/studies MP
- FDA controls MP imports
- FDA reviews submissions to market a new MP that is like a currently marketed MP
- FDA reviews submissions to market a new novel MP
- FDA orders or conducts a post-marketing study
- FDA reviews reports
- FDA investigates outbreaks
- FDA reviews labelling

Data that could help FDA answer the questions

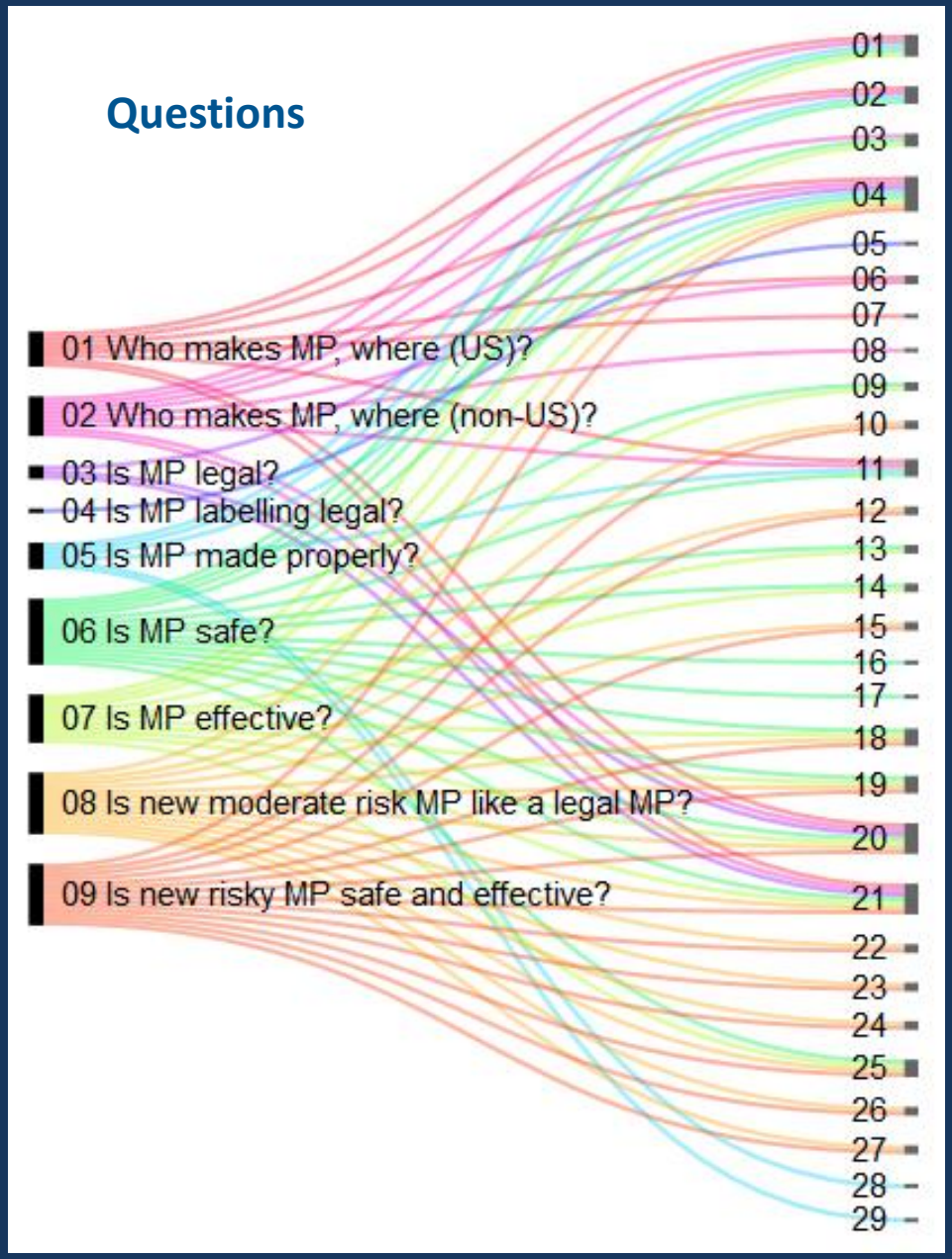
1. MP supply chain
2. Manufacturer credit records
3. Global trade data
4. MP sales, commerce records
5. MP advertising, any media
6. Crowdsourced data about manufacturers
7. State/local records of businesses
8. Registration and listing from foreign governments
9. FDA registration and listing
10. Post-marketing site inspection data
11. Drone-collected data at site inspection
12. Marketed MP inspection data
13. Photos of MP
14. Spectral data acquired regarding MP
15. Import records

Data that could help FDA answer the questions

16. NGS data from outbreak-related specimens
17. NGS data from specimens collected at a US port
18. NGS-based in-vitro diagnostic device: data
19. NGS-based in-vitro diagnostic device: code
20. MP maintenance records
21. MP use records
22. Pre-clinical data collected for MP development
23. Clinical data collected for MP development
24. Data unreleased by manufacturer
25. Healthcare data
26. Post-marketing reports
27. Published scientific reports
28. Good Manufacturing Practices guidance
29. Manufacturing process in premarket submissions

HPC is useful or necessary to use these data

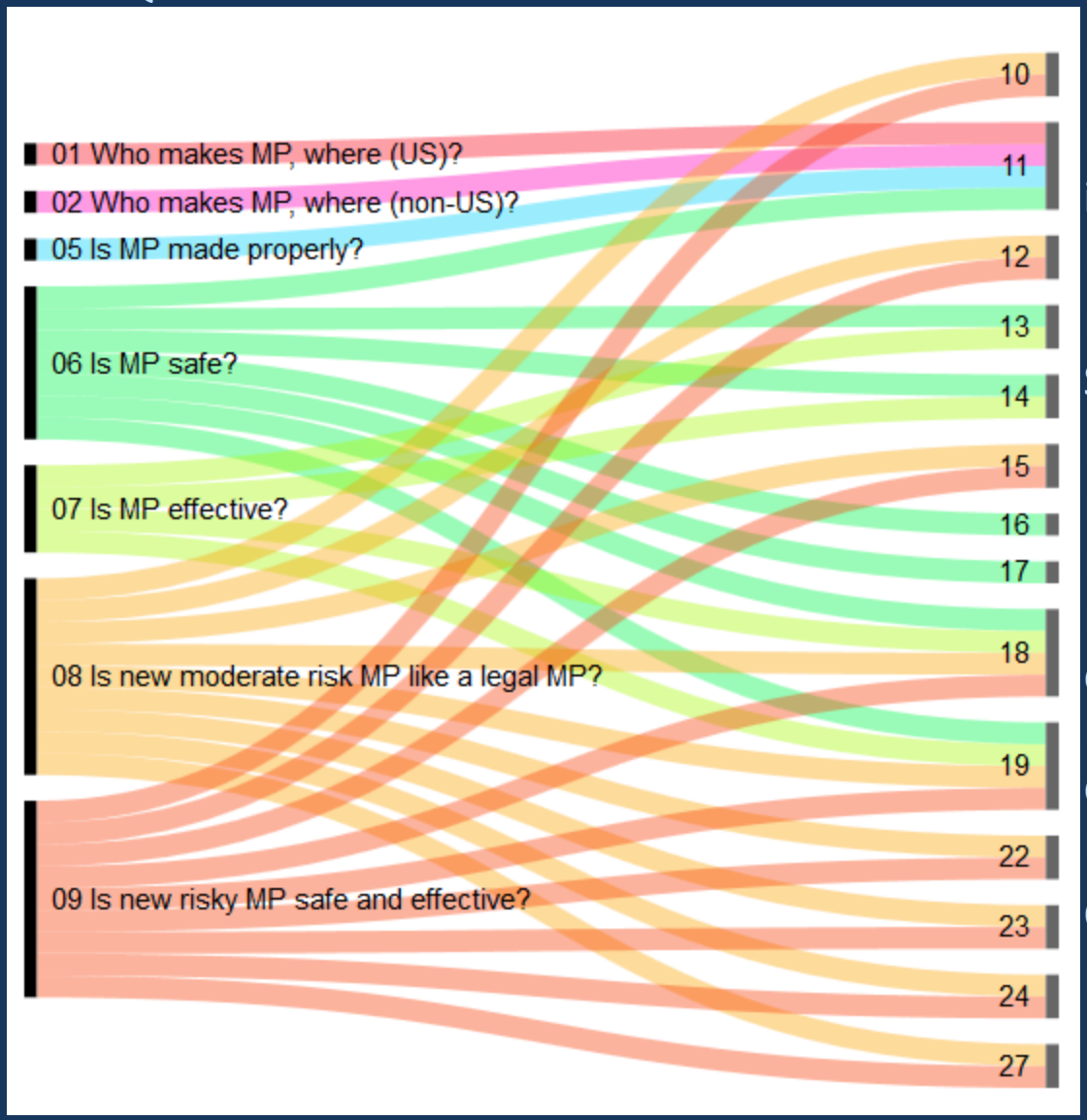
Links between the questions and all the listed data



All data types

1. MP supply chain
2. Manufacturer credit records
3. Global trade data
4. MP sales, commerce records
5. MP advertising, any media
6. Crowdsourced data about manufacturers
7. State/local records of businesses
8. Registration and listing from foreign governments
9. FDA registration and listing
10. Post-marketing site inspection data
11. Drone-collected data at site inspection
12. Marketed MP inspection data
13. Photos of MP
14. Spectral data acquired regarding MP
15. Import records
16. NGS data from outbreak-related specimens
17. NGS data from specimens collected at a US port
18. NGS-based in-vitro diagnostic device: data
19. NGS-based in-vitro diagnostic device: code
20. MP maintenance records
21. MP use records
22. Pre-clinical data collected for MP development
23. Clinical data collected for MP development
24. Data unreleased by manufacturer
25. Healthcare data
26. Post-marketing reports
27. Published scientific reports
28. Good Manufacturing Practices guidance
29. Manufacturing process in premarket submissions

Links between the questions and physical data

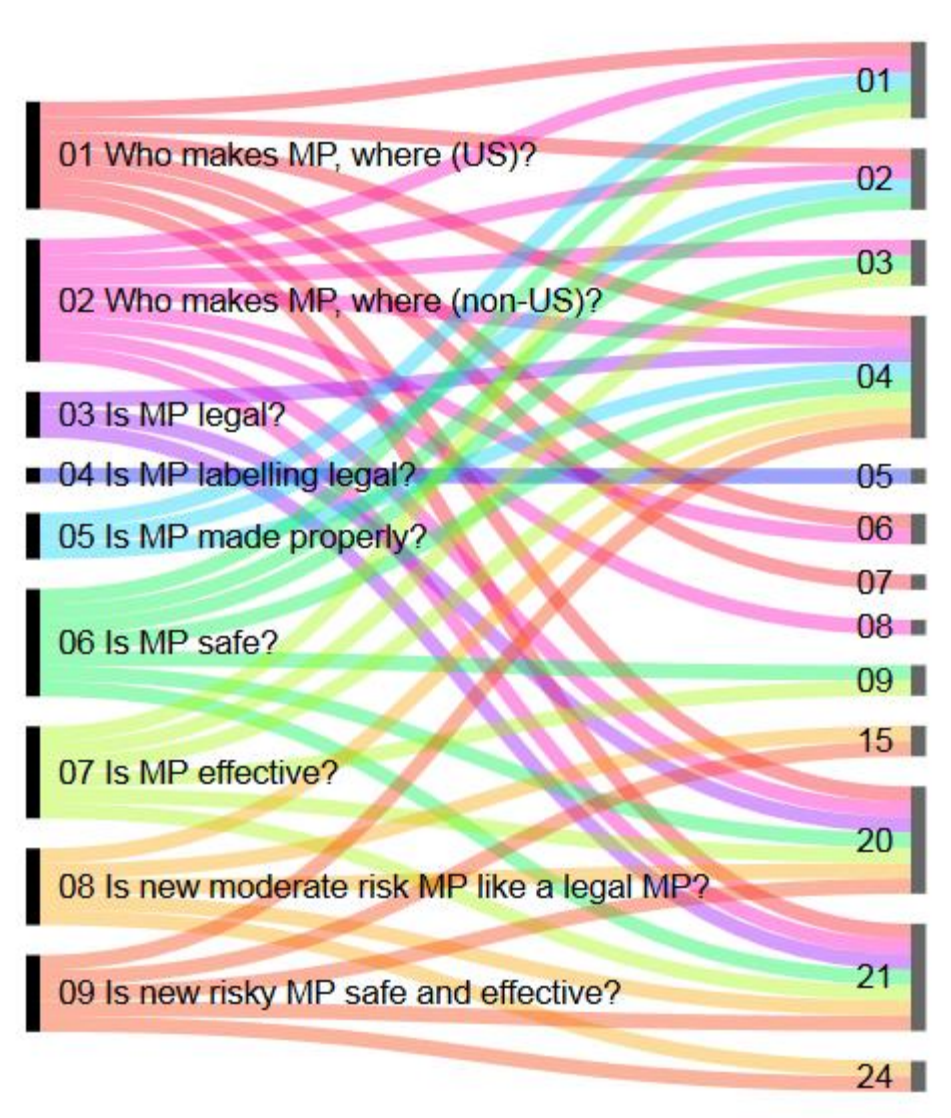


Physical data types

- Post-marketing site inspection data
- Drone-collected data at site inspection
- Marketed MP inspection data
- Photos of MP
- Spectral data acquired regarding MP
- Import records
- NGS data from outbreaks
- NGS data from specimens - US port
- NGS-based in-vitro diagnostic device: data
- NGS-based in-vitro diagnostic device: code
- Pre-clinical data for MP development
- Clinical data for MP development
- Data unreleased by manufacturer
- Published scientific reports

Links between the questions and trade data

Questions



Trade data types

MP supply chain

Manufacturer credit records

Global trade data

MP sales, commerce records

MP advertising, any media

Crowdsourced data about manufacturers

State/local records of businesses

Registration and listing from foreign governments

FDA registration and listing

Import records

MP maintenance records

MP use records

Data unreleased by manufacturer

Links between the questions and clinical data

Questions

Clinical data types

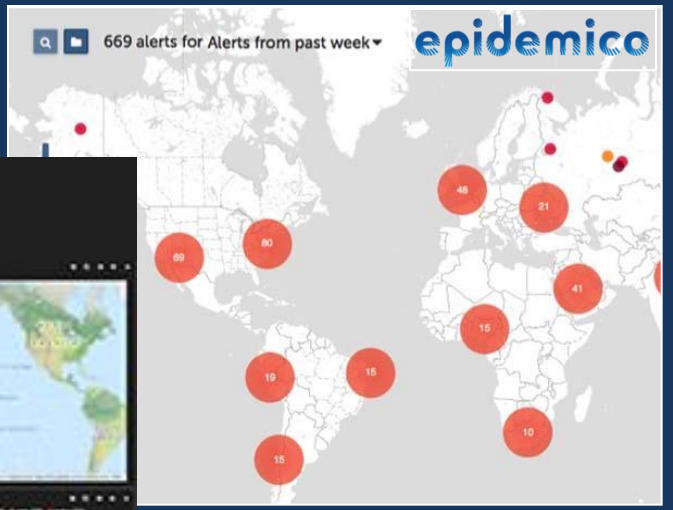
- 04 MP sales, commerce records
- 20 MP maintenance records
- 21 MP use records
- 23 Clinical data collected for MP development
- 24 Data unreleased by manufacturer
- 25 Healthcare data
- 26 Post-marketing reports
- 27 Published scientific reports

State of HPC activities at FDA

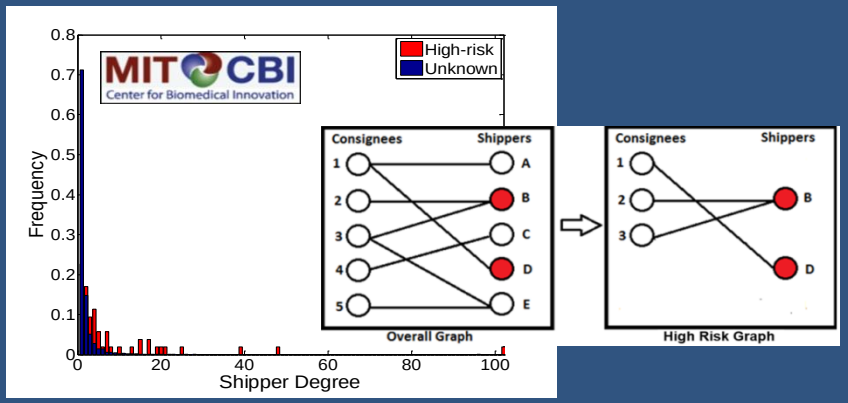
Activity	Current state
China Safety Initiative	Developing dashboards and models to predict unsafe or ineffective imports, using data from trade, Chinese FDA, Chinese media, and FDA
openFDA	• API of large data (two with >4million records each) • >100 processes/sec (designed up to 300/sec)
precisionFDA	In closed beta to develop a cloud platform to exchange/ leverage expertise regarding genomics tests
GenomeTrakr and Chillax	Has shortened time to resolve an outbreak from 12 days to a few hours (can scale to 100Ks nodes).
Sentinel Initiative	Restructuring electronic health records from very big healthcare providers into standardized and easily usable distributed databases for over 175m US population

China Safety Initiative: Global Landscape Analysis

Weather/Environment signals



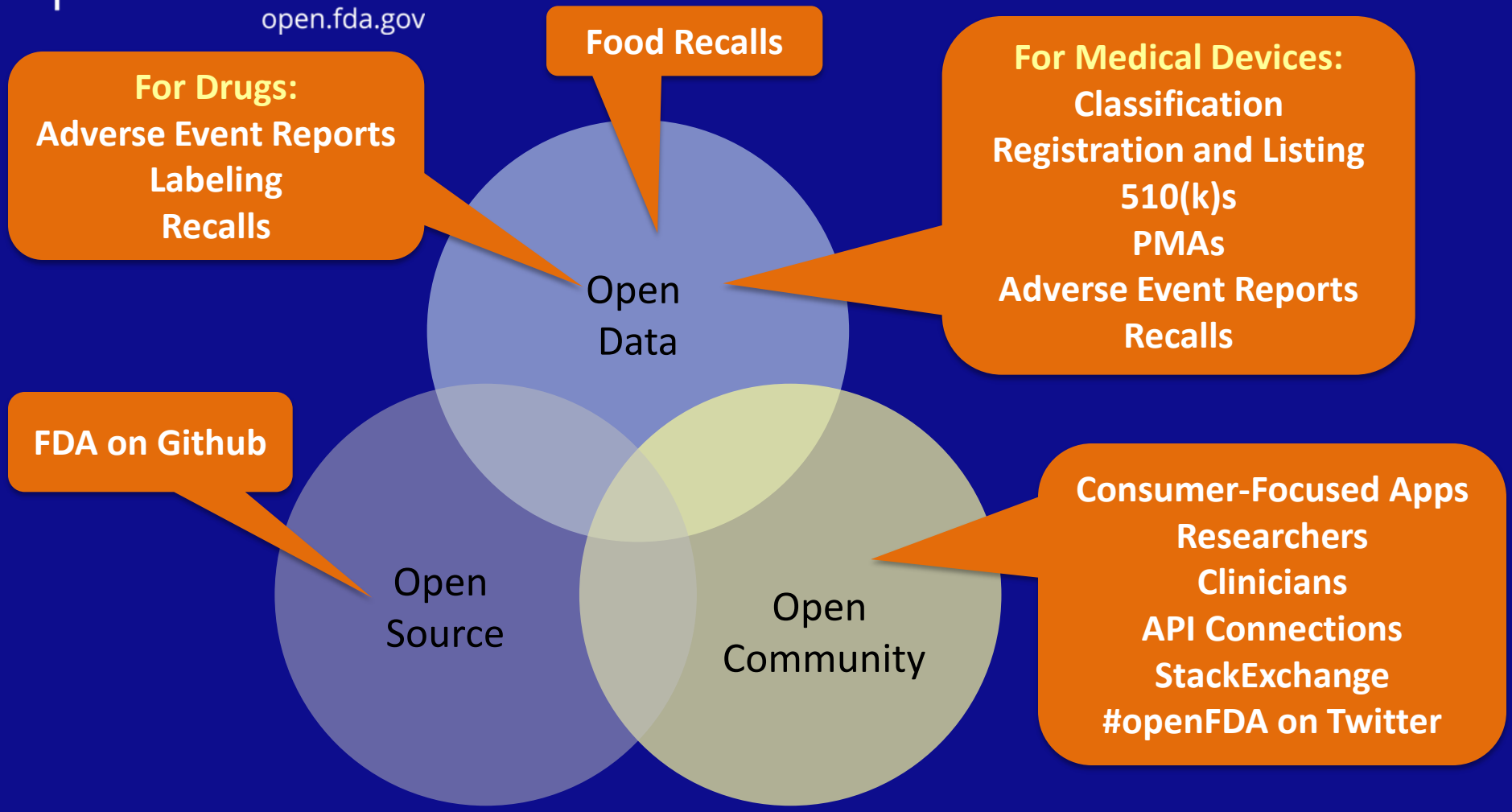
Open Source / Social Media Alerts



Risk Driver/Supply Chain Analysis

Three Components of Transparency and Visibility

openFDA
open.fda.gov



OpenFDA promotes data sharing, data access, and transparency in our regulatory and safety processes, and spurs innovative ideas for promoting public health.

OpenFDA Datasets

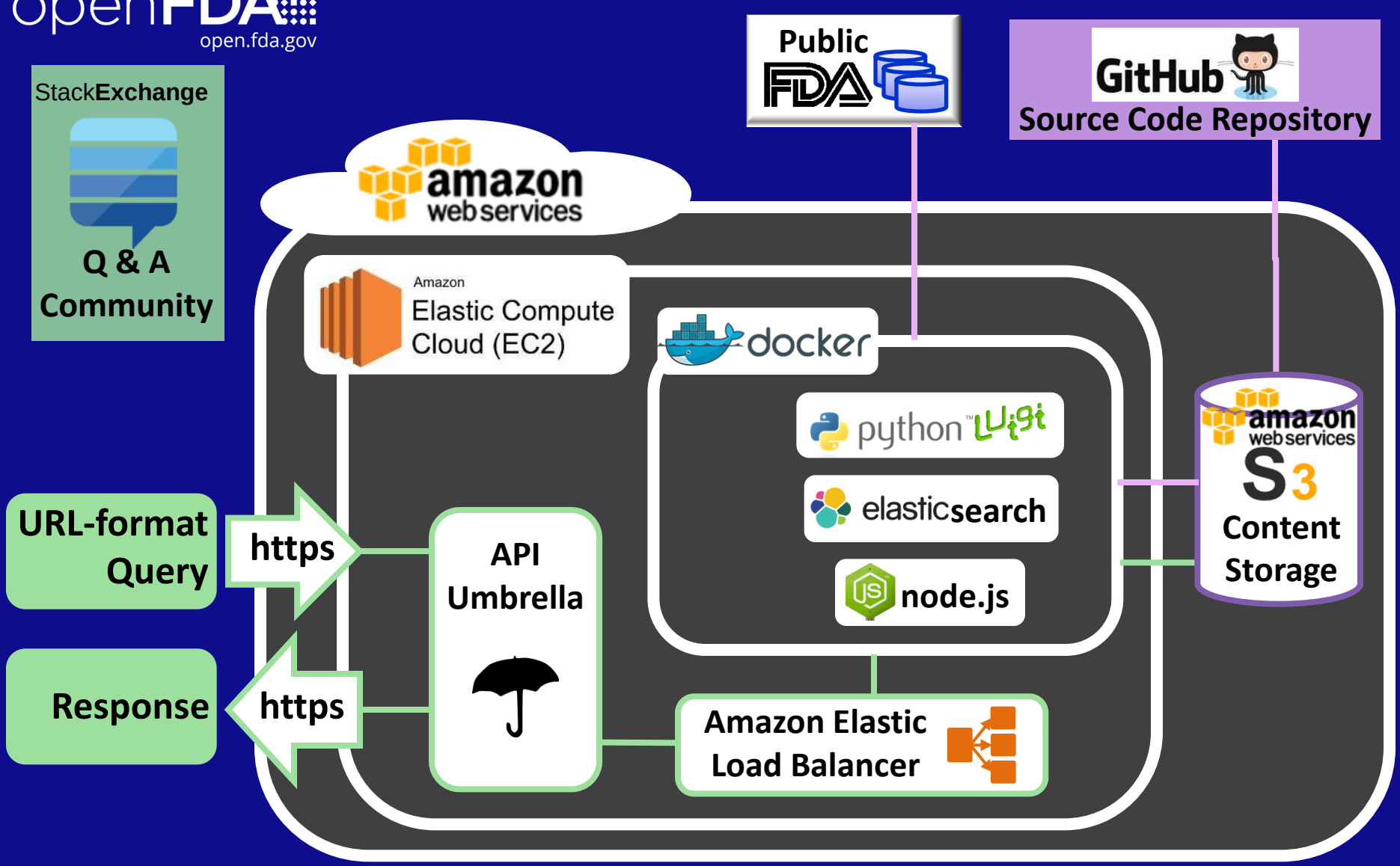
openFDA
open.fda.gov

Unique APIs were developed in FY 2014-15 for readily available datasets (#s as of Aug 2015):

Product	Data	Timeframe	# of records
Drug	Labeling	Current	55K
	Adverse event reports	Since 2003	4.9M
	Recalls	Since 2012	4K
Medical device	Classification	Current	6K
	Registration & Listing	Current	24K establishments >100K devices
	510(k)s (including de novos)	Since 1976	141K
	PMAs (including supplements)	Since 1977	30K
	Adverse event reports	Since 1991	4.2M
	Recalls	Since 2002	9.5K
Food	Recalls	Since 2012	8.5K

Open Source Logical Architecture

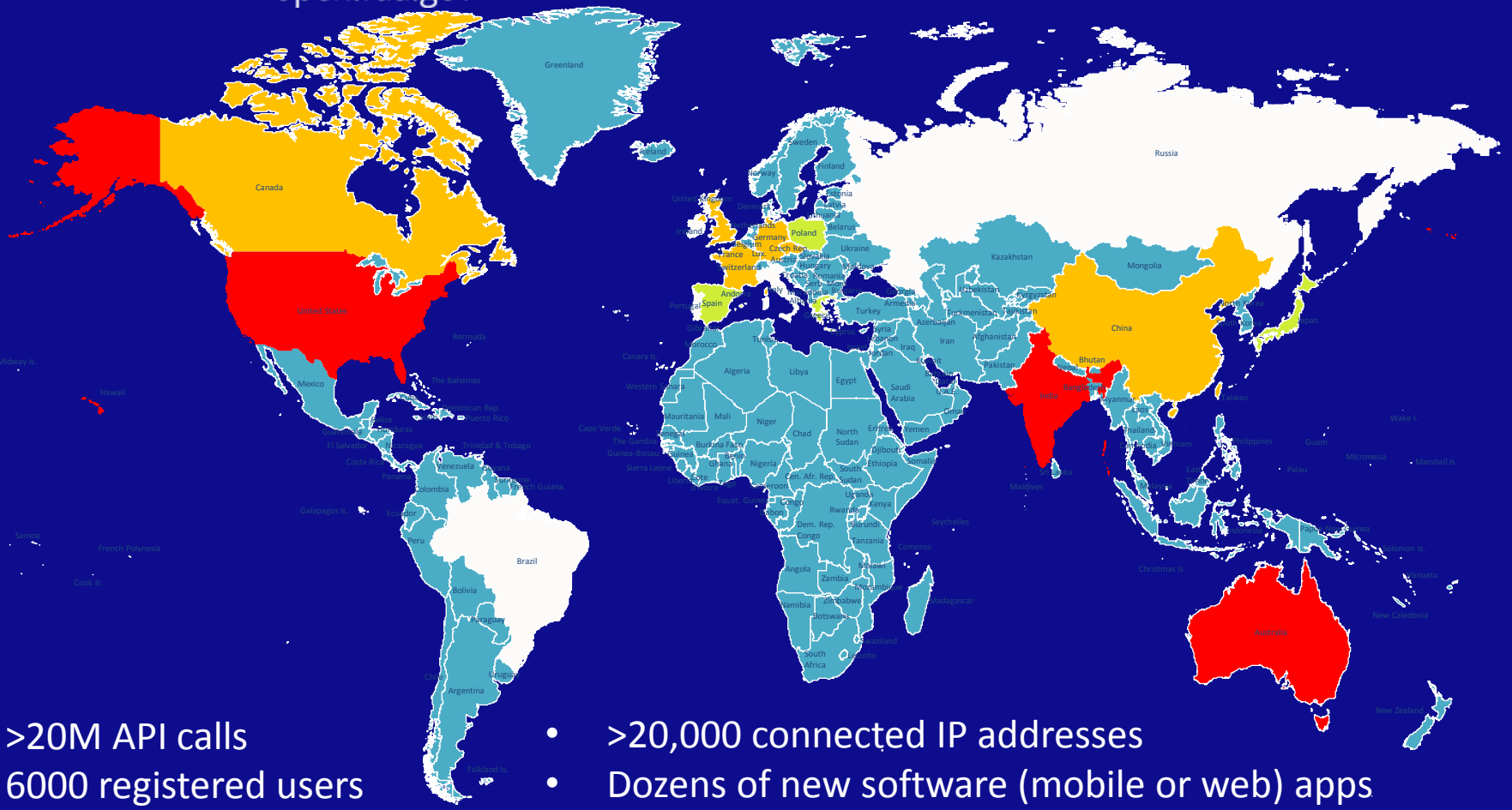
openFDA
open.fda.gov



Data portable to other software. Data/ software portable to other platforms.

Since going live on June 2, 2014

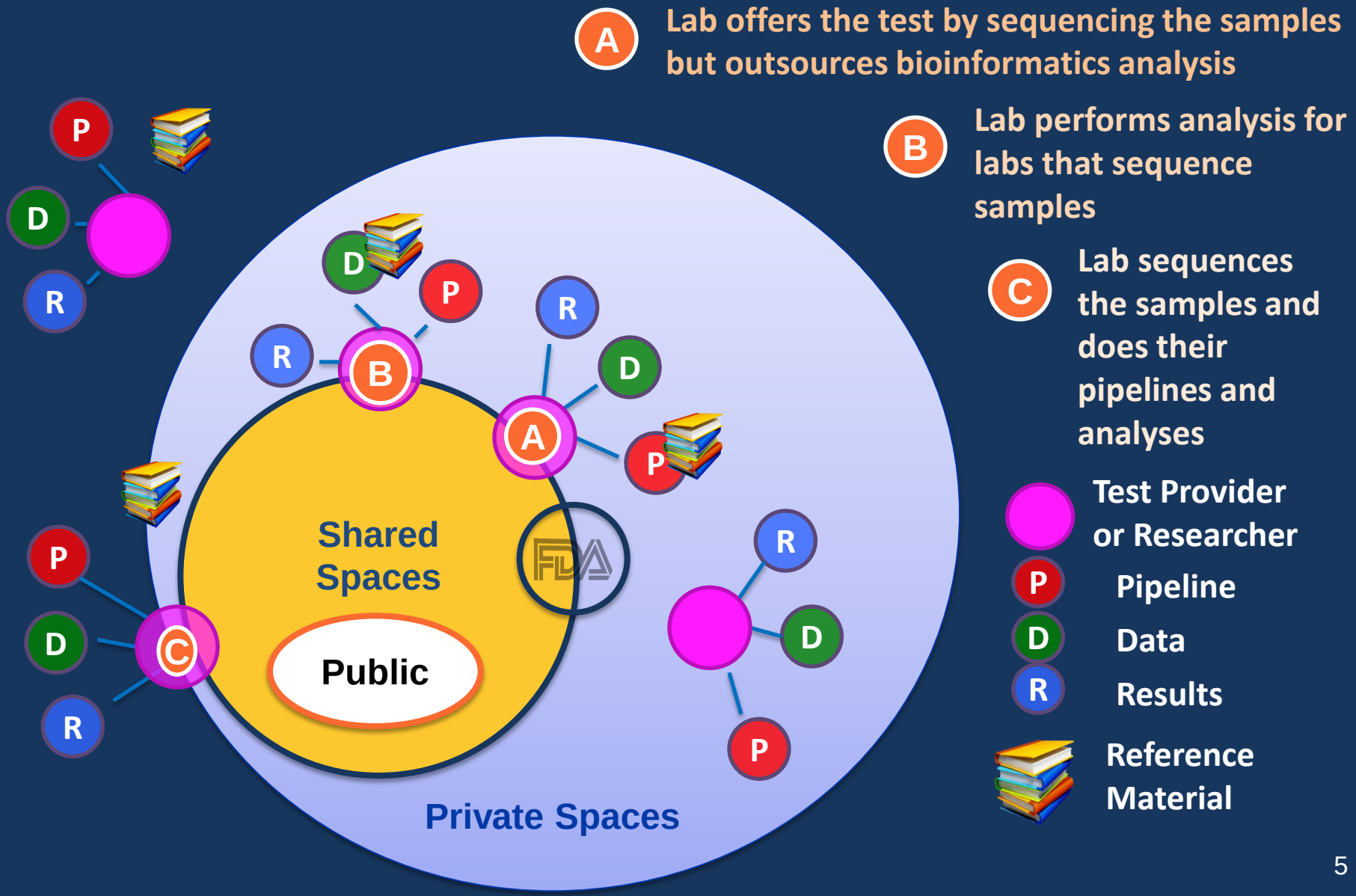
openFDA
open.fda.gov



- >20M API calls
- 6000 registered users
- >20,000 connected IP addresses
- Dozens of new software (mobile or web) apps

>½ of all API calls were from outside the US

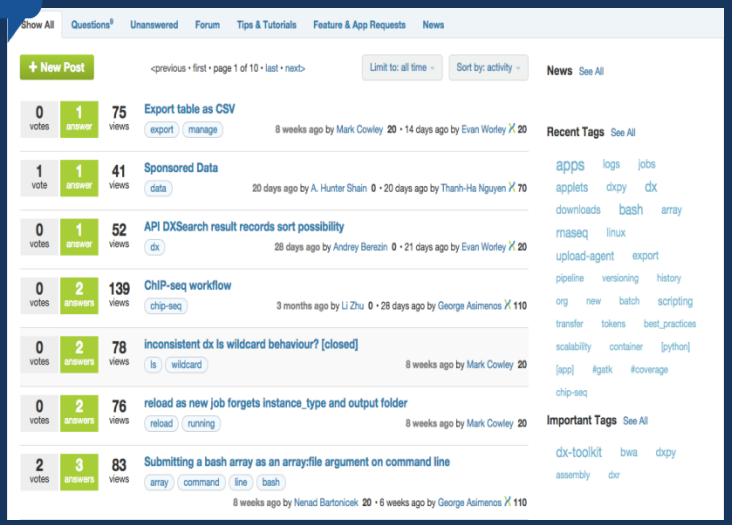
Precision Medicine: Eco-System



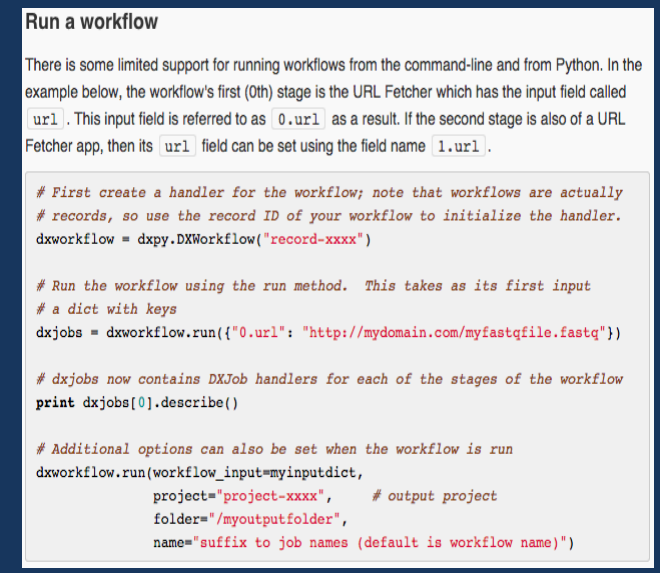
PrecisionFDA: Community Resources

The main goal is to build a strong (*and self-correcting*) community…

Community Wiki (Q&A forum)



Open Source Repository

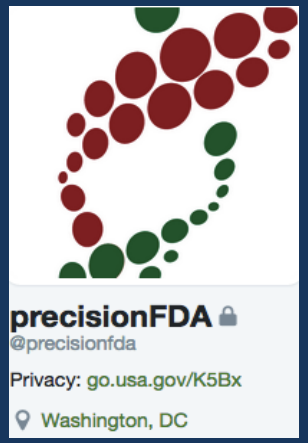


Run a workflow

There is some limited support for running workflows from the command-line and from Python. In the example below, the workflow's first (0th) stage is the URL Fetcher which has the input field called `url`. This input field is referred to as `0.url` as a result. If the second stage is also of a URL Fetcher app, then its `url` field can be set using the field name `1.url`.

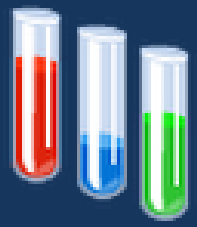
```
# First create a handler for the workflow; note that workflows are actually
# records, so use the record ID of your workflow to initialize the handler.
dxworkflow = dxpy.DXWorkflow("record-xxxx")

# Run the workflow using the run method.  This takes as its first input
# a dict with keys
dxjobs = dxworkflow.run({"0.url": "http://mydomain.com/myfastqfile.fastq"})

# dxjobs now contains DXJob handlers for each of the stages of the workflow
print dxjobs[0].describe()

# Additional options can also be set when the workflow is run
dxworkflow.run(workflow_input=myinputdict,
               project="project-xxxx",     # output project
               folder="/myoutputfolder",
               name="suffix to job names (default is workflow name)")
```

Social Media



GenomeTrakr

Infectious outbreak samples



External health labs and 10 FDA field labs

Pathogens' genome sequences

Sample_ID	I7_Index_	I5_Index_	Sample_Project
SAMN02345037	N702	N503	PRJNA186035
SAMN02345033	N704	N503	PRJNA186035
SAMN02345036	N701	N503	PRJNA186035
SAMN02345035	N702	N504	PRJNA186035
SAMN02345029	N704	N504	PRJNA186035

FDA
Center for Food Safety and Applied Nutrition

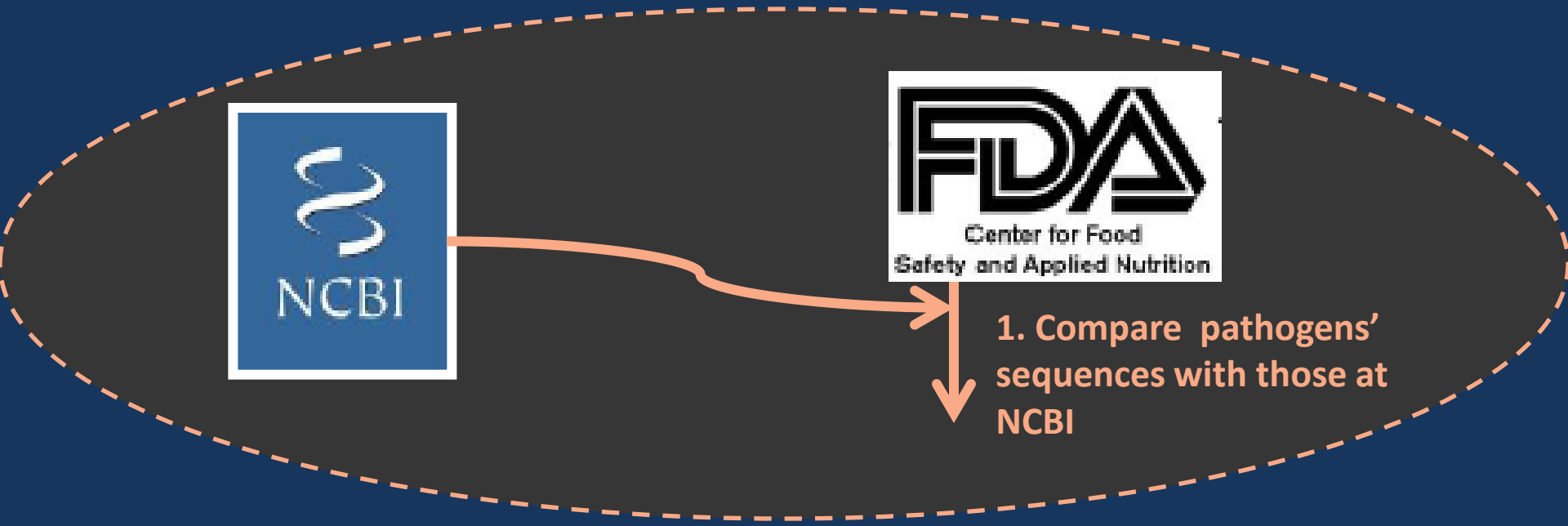
NCBI

1. Compare sequences with those at NCBI

2. Find outbreak root cause

3. Submit new data to NCBI

Chillax



Q: Can the same analysis (#1 above) be done in the cloud?
A: Yes.

Q: Is the same analysis faster in the cloud than in the CFSAN HPC?
A: Yes. ~ 1/18th the time. Cost of $0.5 to $1.0 per isolate.

Sentinel Initiative

- Response to 2007 FDA Amendments Act (FDAAA) mandate:
  - establish a system for active surveillance of drugs
  - use electronic data from healthcare data holders
- Goal: build and implement a new active surveillance system that will eventually be used to monitor all FDA-regulated products.

Sentinel Initiative

- Pilot: Mini-Sentinel
- FDA provides funds and makes decisions
  - Academic partners: 15
    - Planning, operations
    - Methods, protocol development
  - Data partners: 18, with <175M patients
- Has completed dozens of assessments on:
  - Exposures to medical products
  - Diagnoses and procedures
  - Outcomes among those exposed to medical products
  - Impact of FDA actions

What would it take to enhance a current FDA activity with an HPC-enabled activity?

- Scientific evidence that HPC-enabled activity is at least as good (in terms of conclusions and both types of errors) as old activity
- Emotional comfort with “black box” computations
- Focus on a few (1-3) win-win use cases to demonstrate HPC utility

Appendix C

Toward Predictive Biology

with Extreme Computing

Workshop on Future Role of High Performance Computing in Medical Product Decision-making

HIGH PERFORMANCE COMPUTING
INNOVATION CENTER

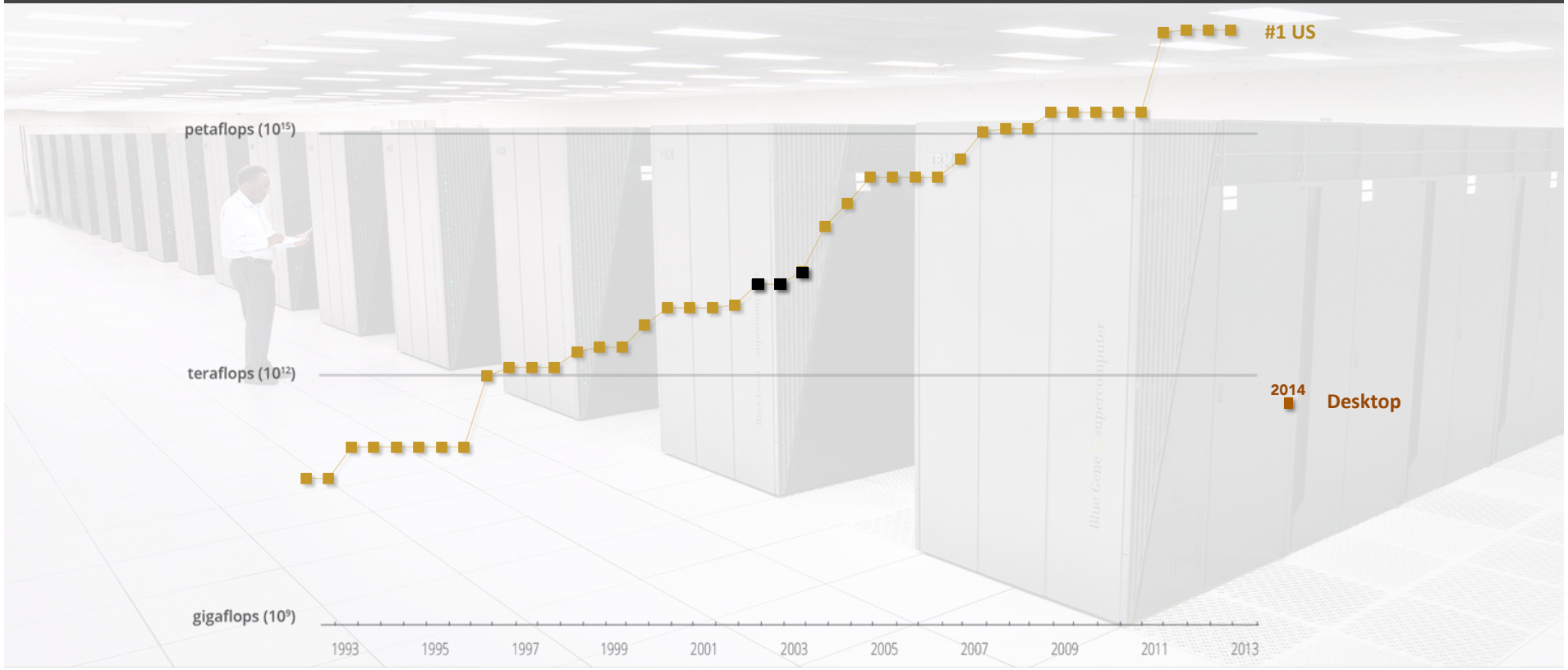
Dr. Frederick Streitz, Ph.D.
Chief Computational Scientist
Director, HPC Innovation Center

Lawrence Livermore National Laboratory

September 10, 2015
Silver Springs, Maryland

LLNL-PRES-668884
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.
Lawrence Livermore National Security, LLC

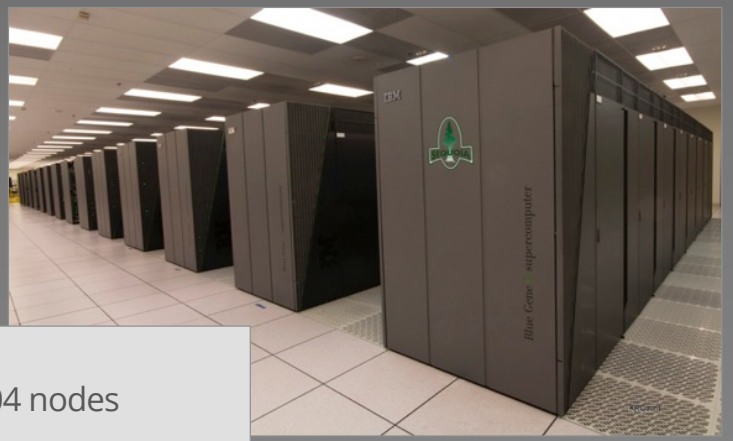
Dept. of Energy leads the **world** in computing



Modern supercomputers are five orders of magnitude more powerful than desktop computers

"Titan" (Cray XK7 ORNL)

- 18,688 nodes
- **299,088 cores**
- **18,688 GPUs**
- 27 Pflop/s peak
- 0.7 + 0.1 PB memory
- 8.2 Megawatts



- 98,304 nodes
- **1,572,864 cores**
- **0 GPUs**
- 20 PFlop/s peak
- 1.6 PB memory
- 9.6 Megawatts

"Sequoia" (IBM BG/Q LLNL)

Enables simulation 100,000X more **realistic** than possible on a desktop

Not 100,000 times **faster?**


Parallel Overhead

Parallel Overhead

100,000 times more powerful is **not** 100,000X faster!

100,000 times more **powerful** can
mean 100,000x more **realistic**

What does **100,000x more realistic** look like?

What does 100,000x more realistic look like?

3.1 megaFLOPS

1980 Video Arcade Game

1x ZiLOG Z80 @ 3.072 MHz

3,100,000

What does 100,000x more realistic look like?

260 gigaFLOPS

2011 PC Video Game

GeForce GT530

3,100,000

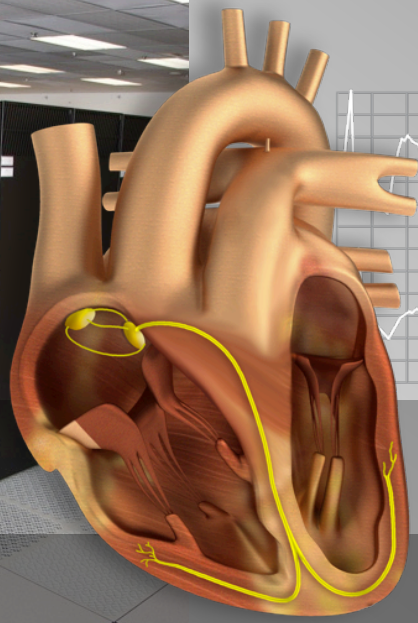
260,000,000,000

Another 100,000x:

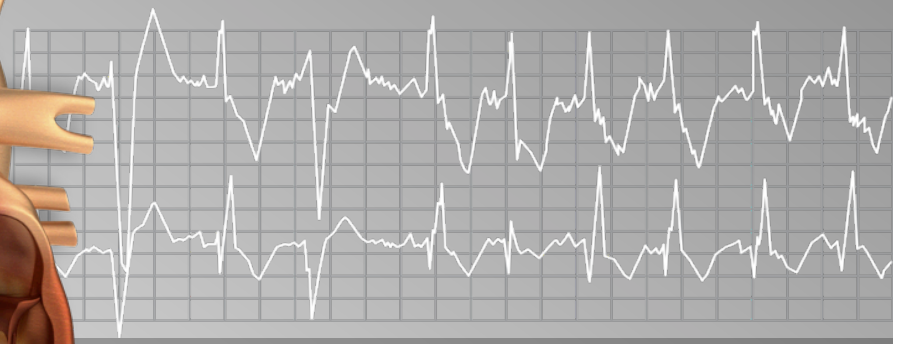
Near-cellular resolution, real-time simulation of a beating human heart

Sequoia computer

20 petaFLOPS



Cardioid: IBM-LLNL Collaboration

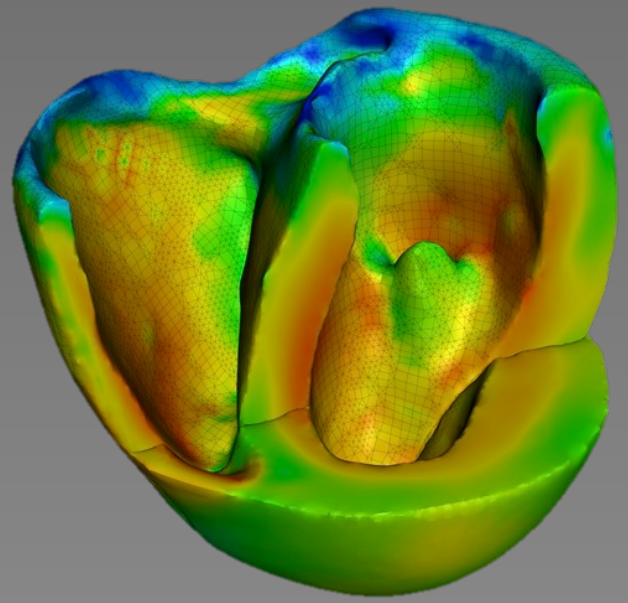


3,100,000

260,000,000,000

20,000,000,000,000,000

Whole heart modeling at cellular resolution in real time



Form multi-disciplinary, multi-institutional team (IBM-LLNL)

Develop high resolution, realistic model of human heart

Create *Cardioid* code to model electrophysiology of heart

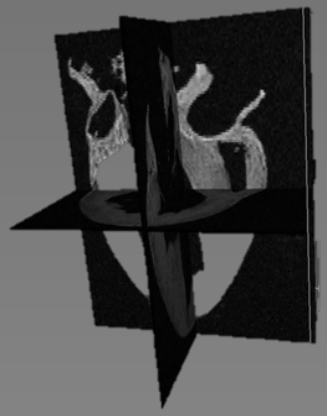
Leverage 20 PF Sequoia (Blue Gene/Q) resource

Investigate development of arrhythmia

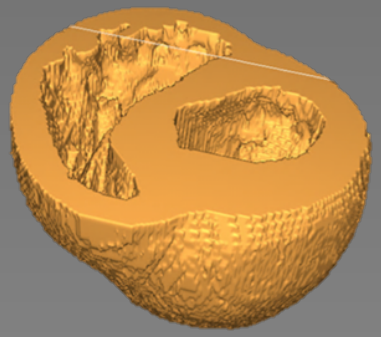
Challenges

- Sudden Cardiac Arrest is a leading cause of death in the U.S. ~ 325,000/year
- Complexity of measured ECG makes identification of mechanisms difficult
- Fast, high-resolution model enables exploration of drug-induced arrhythmia

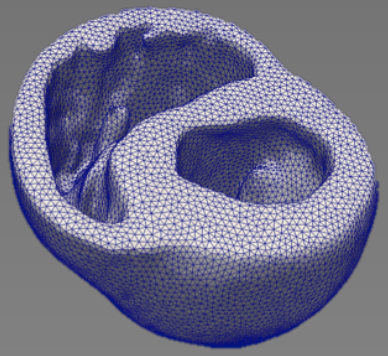
Building a human heart model



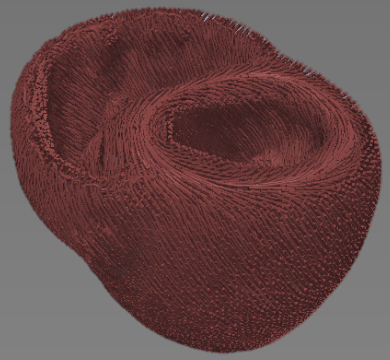
Raw Data



Segmentation



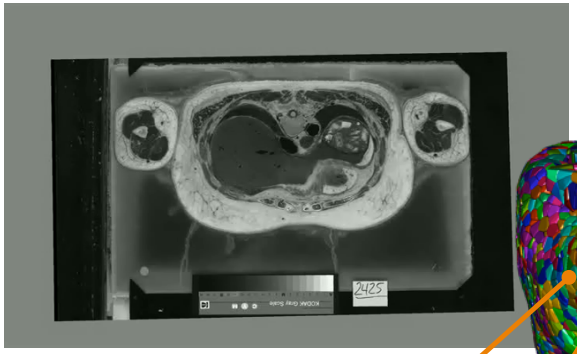
Volume Mesh



Fiber Generation

- Import data from Visible Human Data Set
- Identify and partition ventricular segments
- Mesh the data at sufficient resolution
- Overlay fiber geometry

The Cardioid Model on Sequoia



Anatomical model of human heart

96 racks of Blue Gene/Q in Sequoia

- Complete heart is 370-700 million volume elements in a complex geometry
- Sub-domains of ~200-500 elements must be mapped to each of 1.6 M cores
- Cell model must be computed on every cell for each time step including reaction and diffusion
- Sub-domains must exchange boundary data on each time step
- Code is written from scratch to take advantage of BG/Q hardware
- Extensive algorithmic improvements

Enables simulation of a realistic heart in real-time

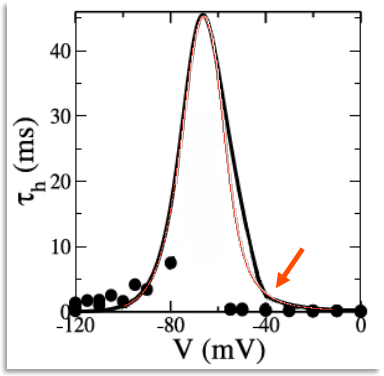
Tuning a heart code to computing hardware

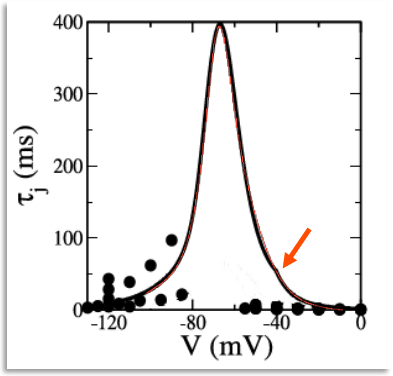
Turning this…

$$\tau^{-1}(V_m) = \frac{1 + e^{(25 - V_m)/10} + 80/(1 + e^{(V_m + 30)/10})}{562 e^{-(V_m + 27)^2/240} + 31}$$

… into this

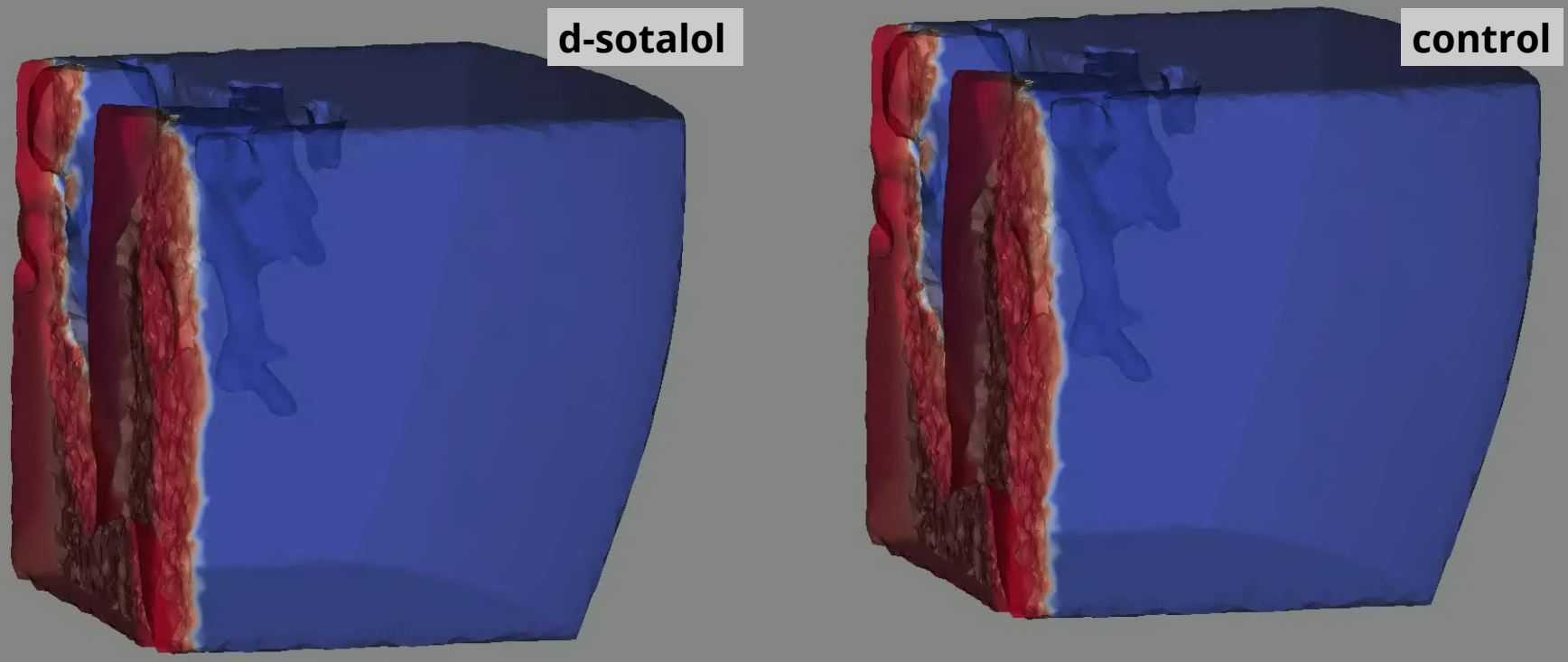
$$f(V_m) \cong \frac{\sum_i a_i V_i^m}{1 + \sum_j b_j V_j^m}$$



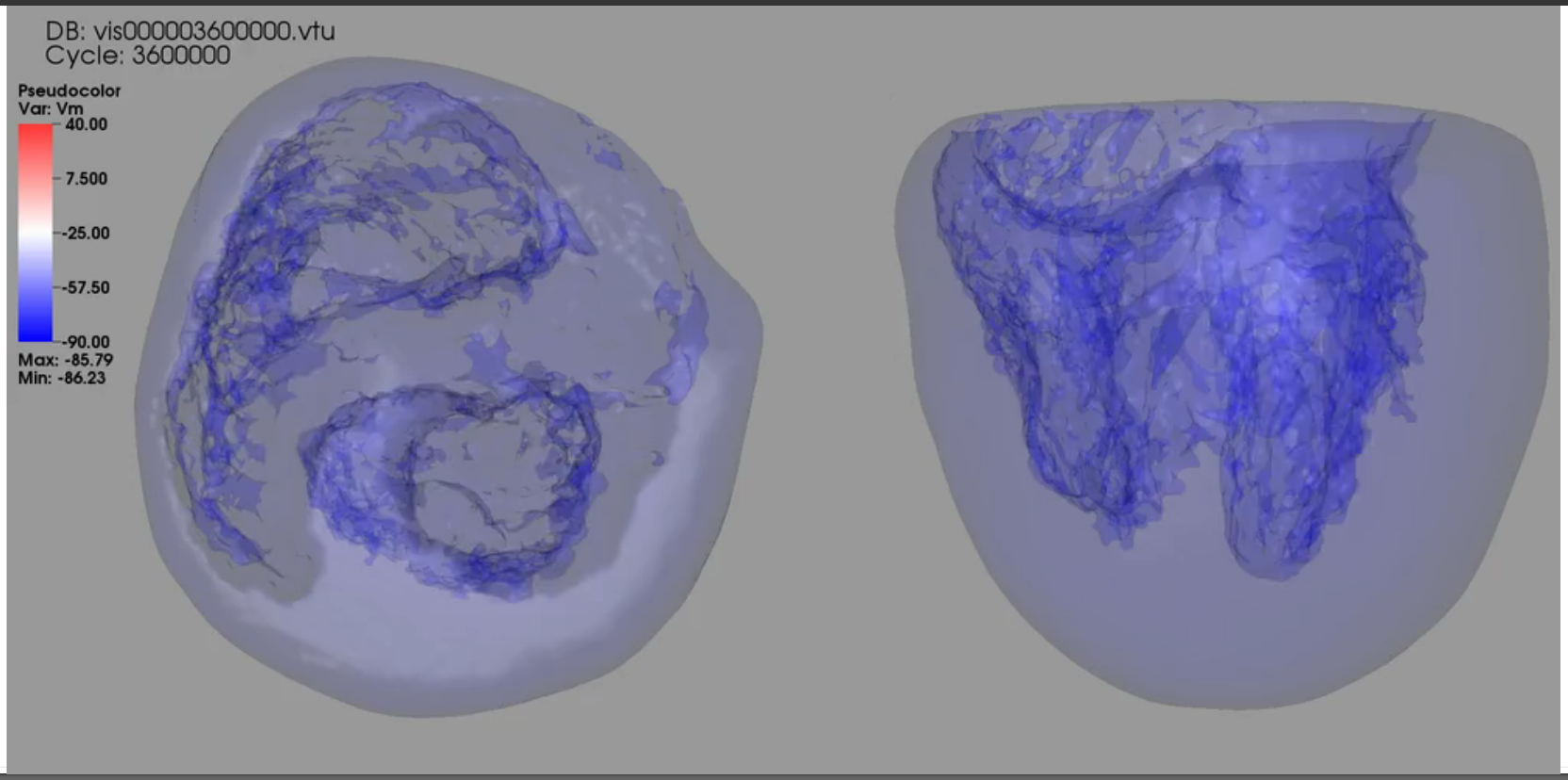


- Two gating time constant functions have discontinuous first derivatives at -40 mV
- Close approximation would require high-order polynomial
- Lower-order fit is sufficient to fit to sparse biological state used to construct the model

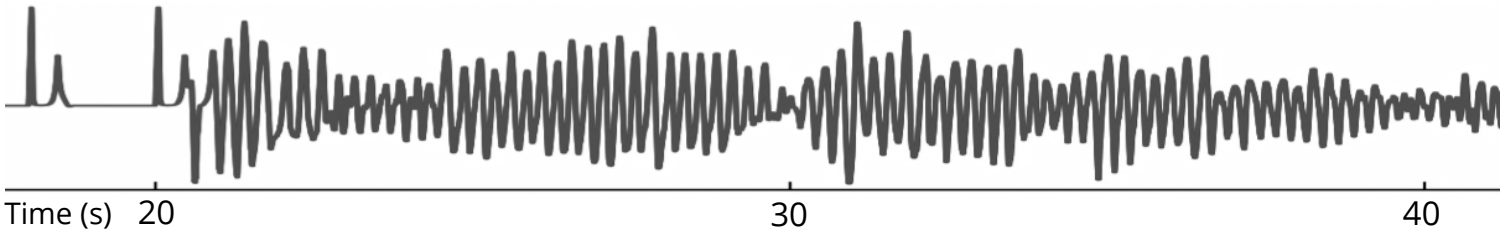
Heart wedge simulation in presence of d-sotalol



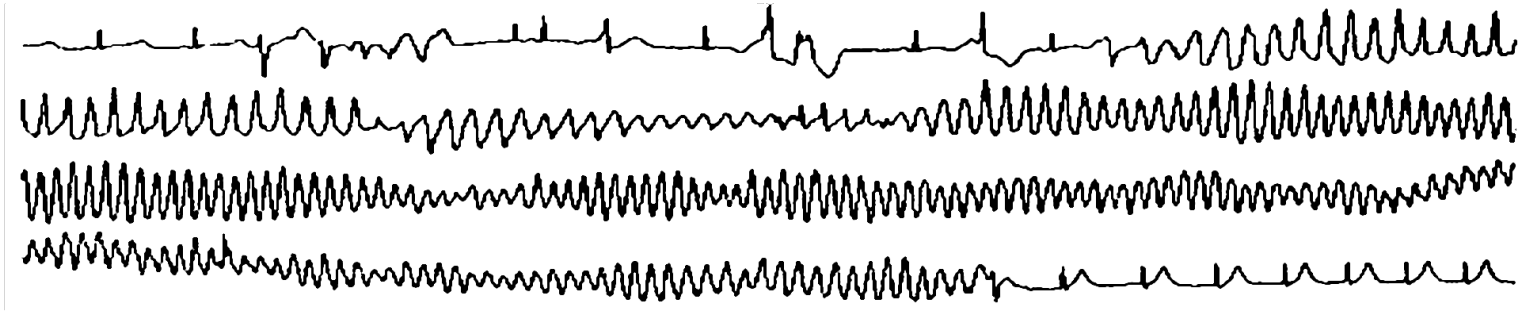
Arrhythmia developing in a whole heart



ECG of transmural re-entrant activation



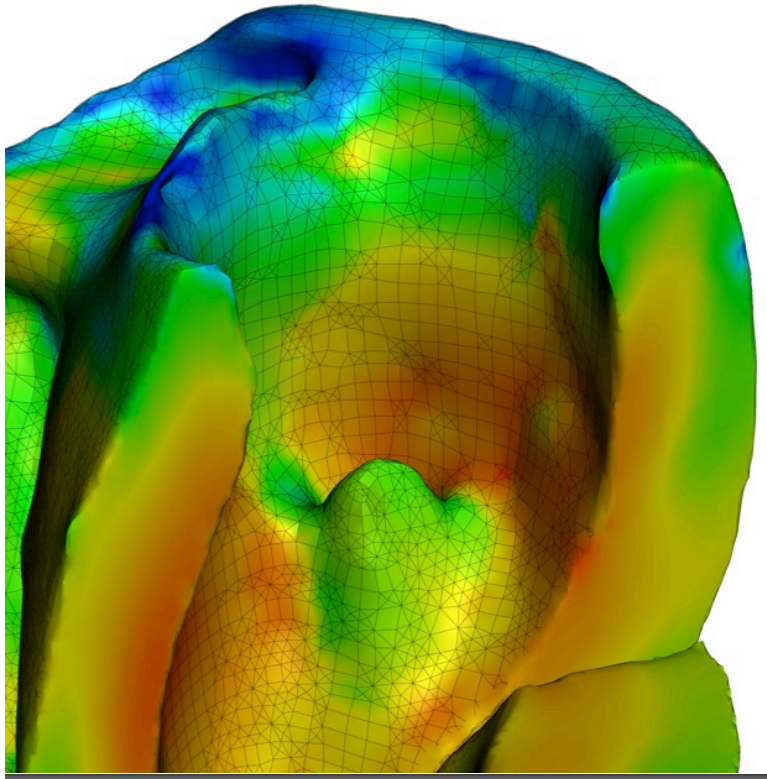
ECG calculated from simulation with E-4031



ECG measured in patient with polymorphic ventricular tachycardia

Kobza R et al. Circulation 2005;111:e173-e174

Potential Applications (Near Term)



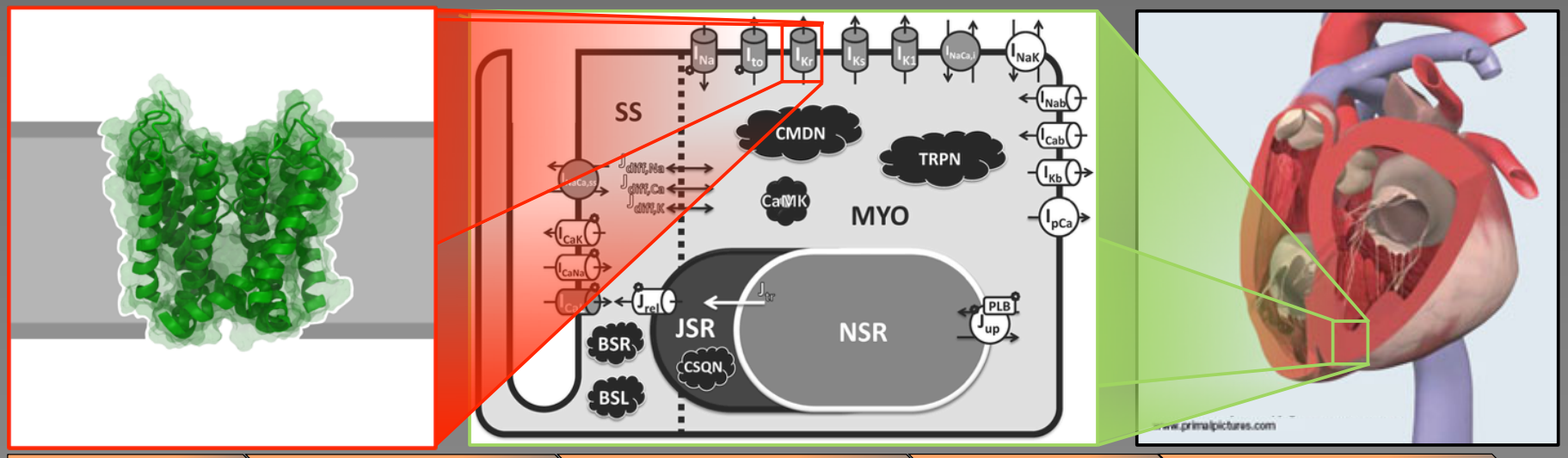
Pharmaceutical

- Drug safety and understanding mechanisms of cardiotoxicity (beyond single-cell effects!)
- New anti-arrhythmics and other cardiovascular drugs
- New drugs for heart failure and metabolic diseases
- Ability to “replay” experiments with different assumptions

Medical devices

- Cardiac Resynchronization Therapy optimization
- Tissue-lead interface models
- New ablation techniques and instruments (including EM effects)
- New treatment approaches of atrial fibrillation

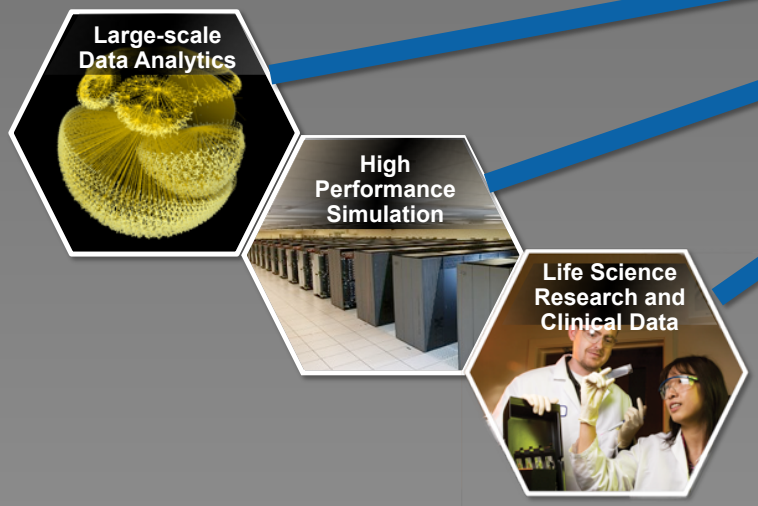
Our goal is a predictive, mechanistic, heart model connecting molecular interaction to clinical outcomes



Develop a near-cellular resolution, real-time simulation of a beating human heart connecting molecular kinetics to clinical outcomes

High performance computing is critical for developing a predictive biology capability to address growing crises in health and biosecurity

- **Rapidly accelerating biosecurity threats**
- **Emerging infectious disease challenges**



- *Today the average time to develop a new drug is ten years, at a cost of $B*
- *Drugs entering clinical trials fail > 90% of the time*

Validated simulations of complex biological systems will be an important tool in responding to future biological threats

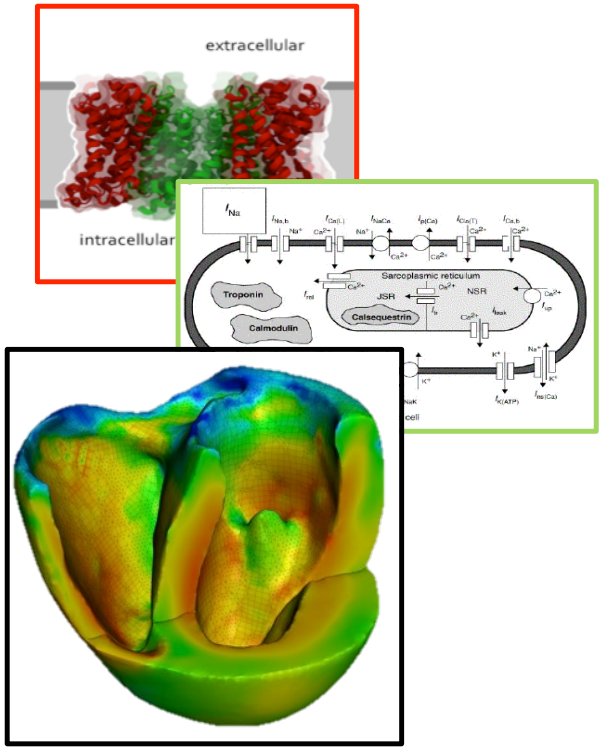
Predictive pharmacology – Simulations of small molecule cardio-toxicity

Cardio-toxicity is the most frequent cause of drug failure. Organ-scale simulations will demonstrate and validate ability of simulation to predict effects.

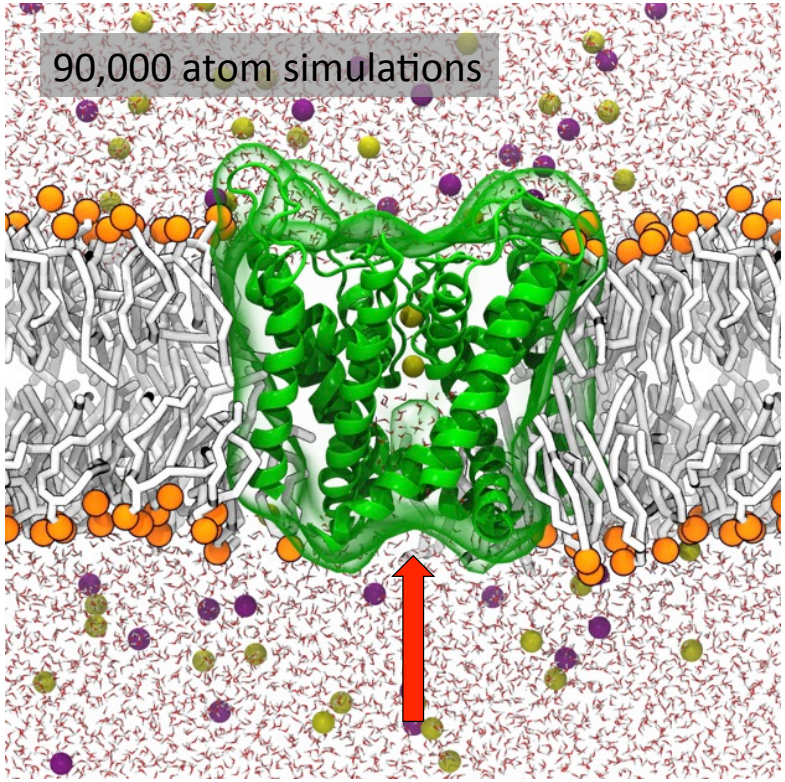
Partnership with Harvard (Loscalzo, Sorger) and UC Davis (Clancy)

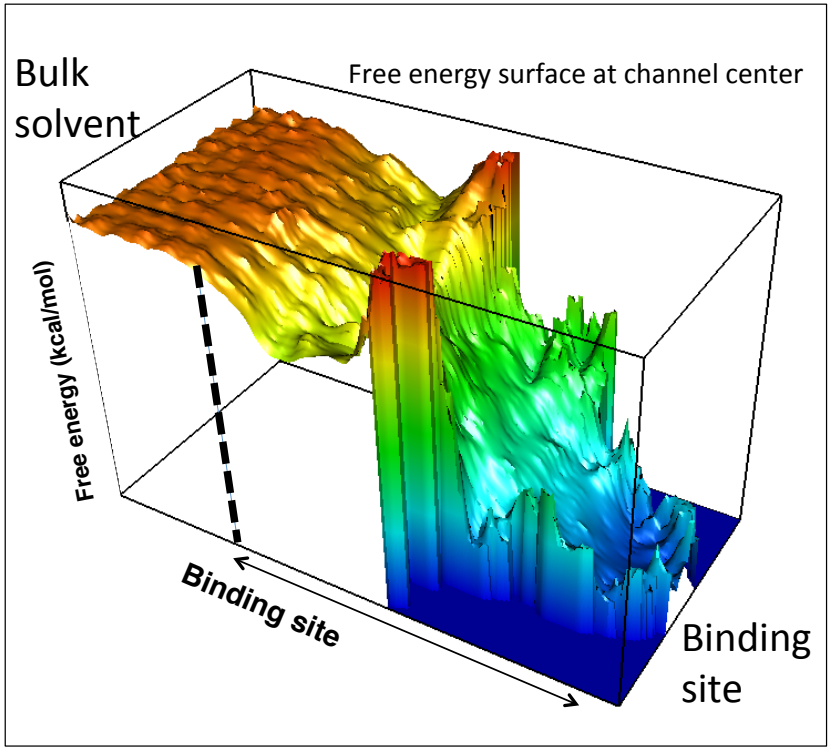
Pilot concept

- Expand Cardioid simulation physiology models
- Demonstrate effect of selected drugs on pathways
- Add mechanistic, atomistic scale



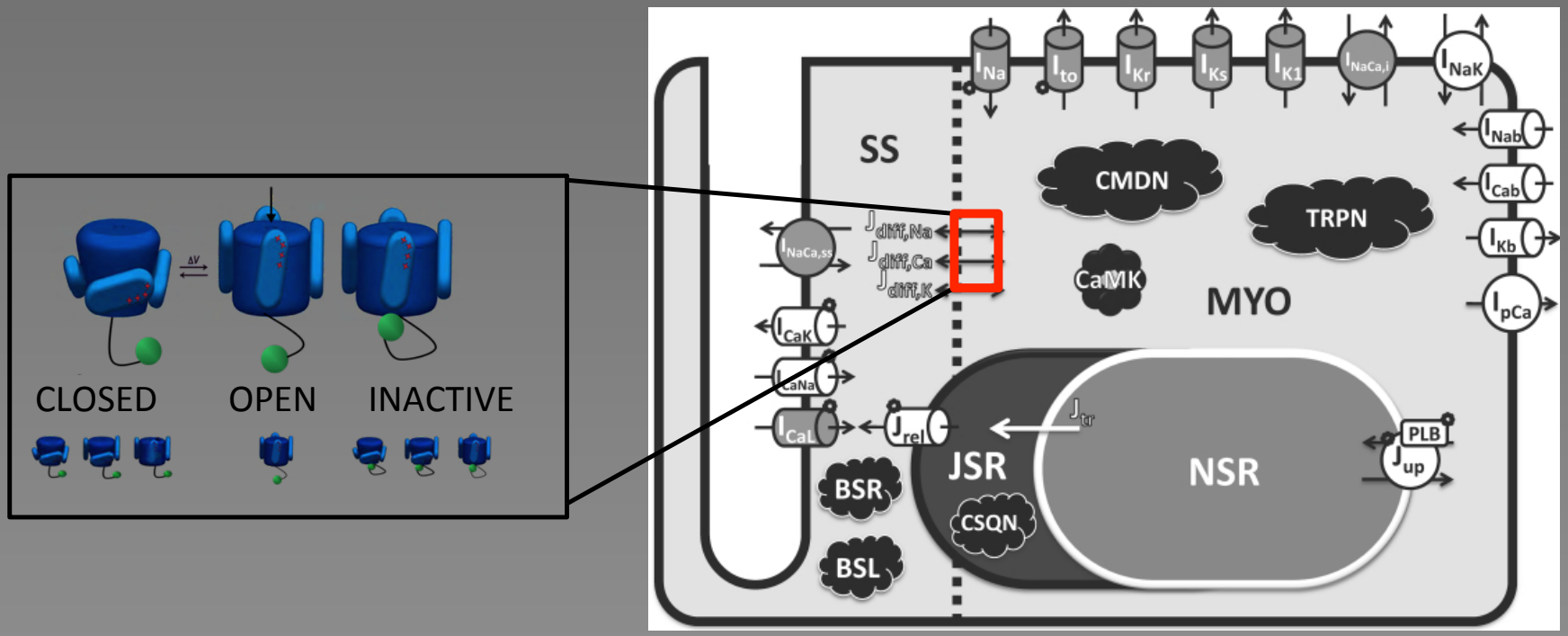
Molecular modeling to determine drug dependent parameters





ΔG = ~ 9.5 kcal/mol K_d ~ 1.3 x 10^{-7} M

O'Hara-Rudy model adds substantial complexity to *Cardioid*



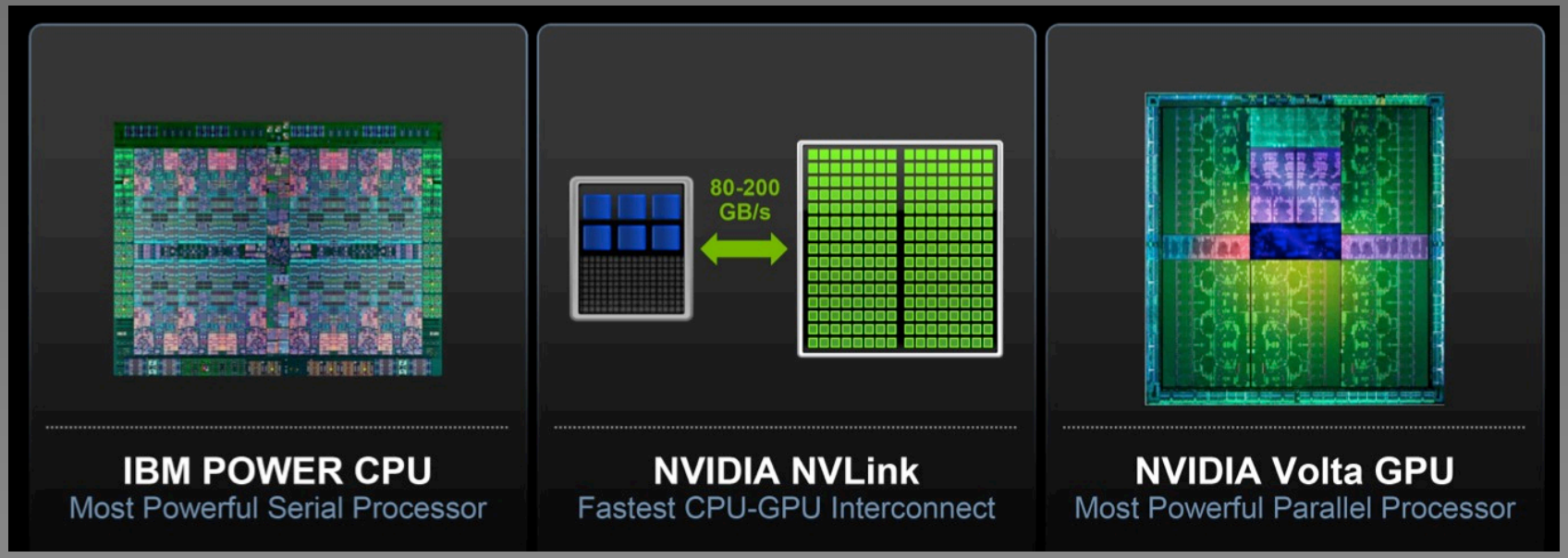
O'Hara-Rudy is formulated from experimentally derived human myocyte data

Next generation computing from IBM & NVIDIA: CORAL

System Feature	Summit (2017)	Titan (2013)
Peak System Performance	> 150 PFlop/s	27 PFlop/s
Peak Node Performance	> 40 TFlop/s	> 1 TFlop/s
# Nodes	>3400	18,688
CPU	IBM Power9	AMD Opteron (0.1 TFlop/s)
GPU	NVIDIA Volta	NVIDIA Keplar (1.3 TFlop/s)
Memory per Node	> 512 GB	> 32 GB
NVRAM per Node	800 GB	0
Node interconnect	NVIDIA NVLink (80-200 GB/s)	PCIe 2 (25 GB/s)
System interconnect	Infiniband EDR (> TB/s)	Cray Gemini (168 GB/s)
Peak Power Consumption	10 MW	8 MW

- Anticipated computing advances move towards us closer to exascale computing
- Architectures are merging data analytics and simulation

Next generation computing from IBM & NVIDIA: 120+ Pflop/s



- Anticipated computing advances move towards us closer to exascale computing
- Architectures are merging data analytics and simulation

Parting thoughts

Factor of 10^4 or 10^5 in computational is transformational

- Renders earlier "impossible" algorithms tractable
- Allows rapid exploration of massive parameter space
- Enables investigation of cooperative effects

Ability to create usefully predictive simulations requires qualified models with believable parameters

- Close collaboration (open communication) between clinicians, lab scientists and computational scientists
- Access to available data and experimental design

Exciting advances in supercomputing capability are on the horizon

- Enables progress toward goal of predictive biology

Acknowledgements

Lawrence Livermore Nat'l Lab

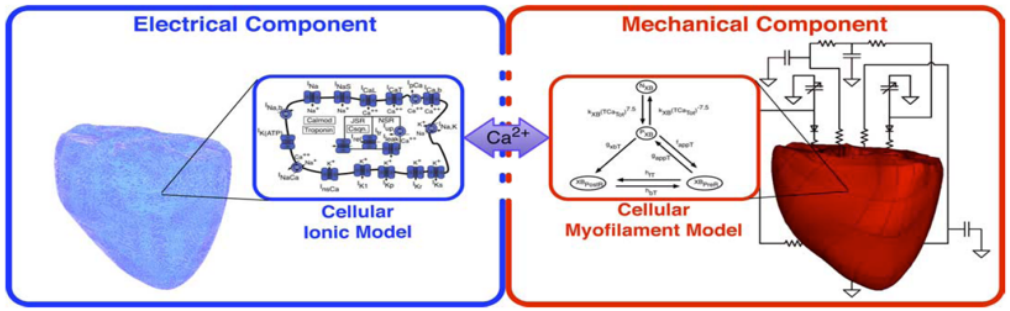
Art Mirin
David Richards
Jim Glosli
Jean-luc Fattebert
Erik Draeger
Tomas Oppelstrup
Bor Chan
Liam Krauss

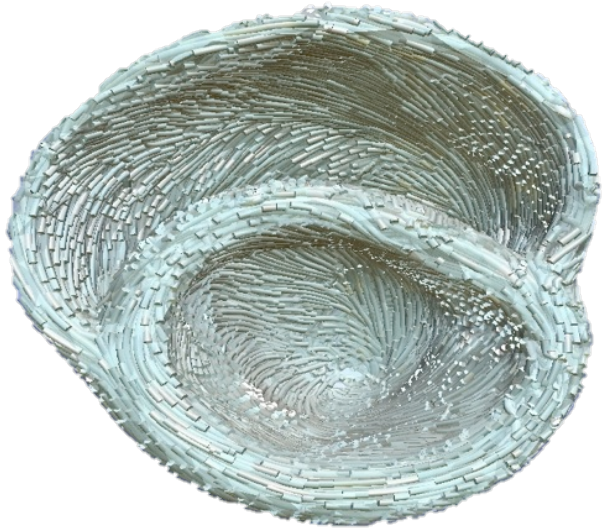
IBM

Jeremy Rice
Slava Gurev
Sophia Wen
John Gunnels
Matthias Reumann
Changhoan Kim
John Magerlein

Electromechanical models of the ventricles

Goal of building predictive model by increasing spatial resolution, mechanistic detail and speed of solutions

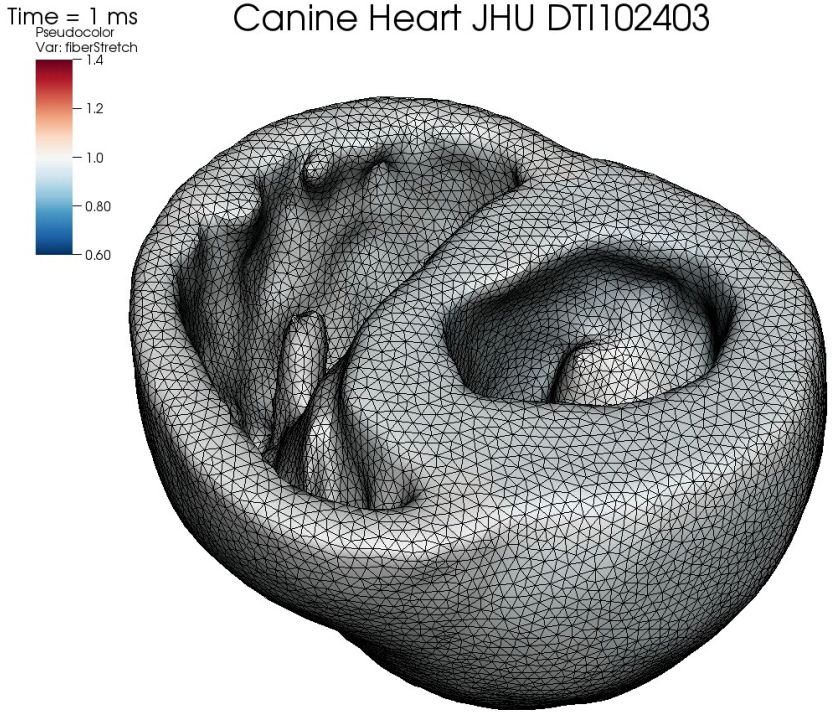




- Electrophysiological model informs electromechanical model through Ca2+ current
- Reconstruction of fiber geometry in tetrahedral model based on DTMRI
- Solve motion of incompressible material using finite element approach

Helm et al. 2005
Constantino et al. 2010

Preliminary view: high-res beating heart

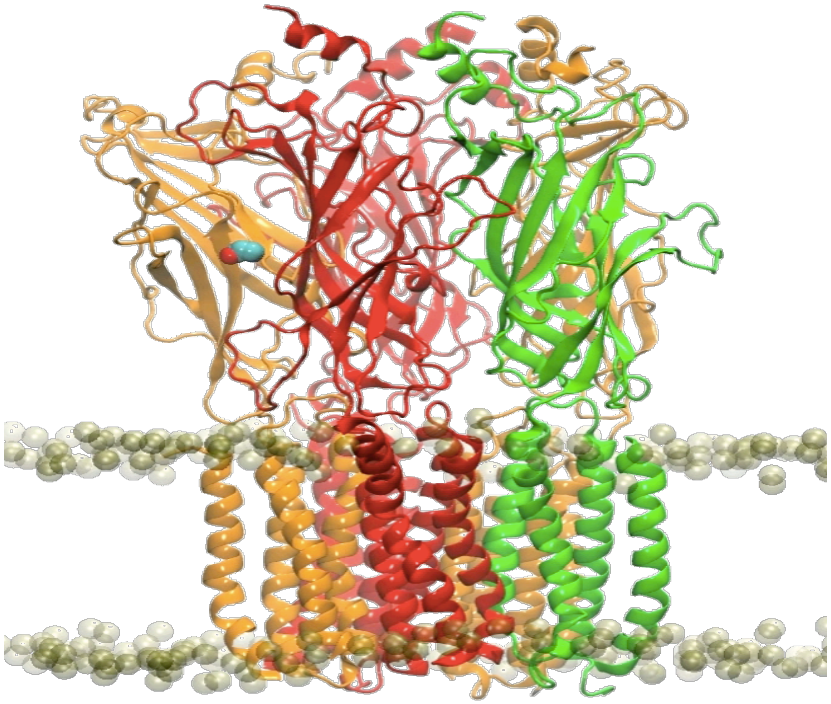


- Contraction in ventricles coupled with electrophysiology
- Not yet real time
- Enables modeling of fluid flow in normal and diseased hearts
- Working to add atrial geometry and valves

Molecular methods can provide fundamental understanding of ligand interactions

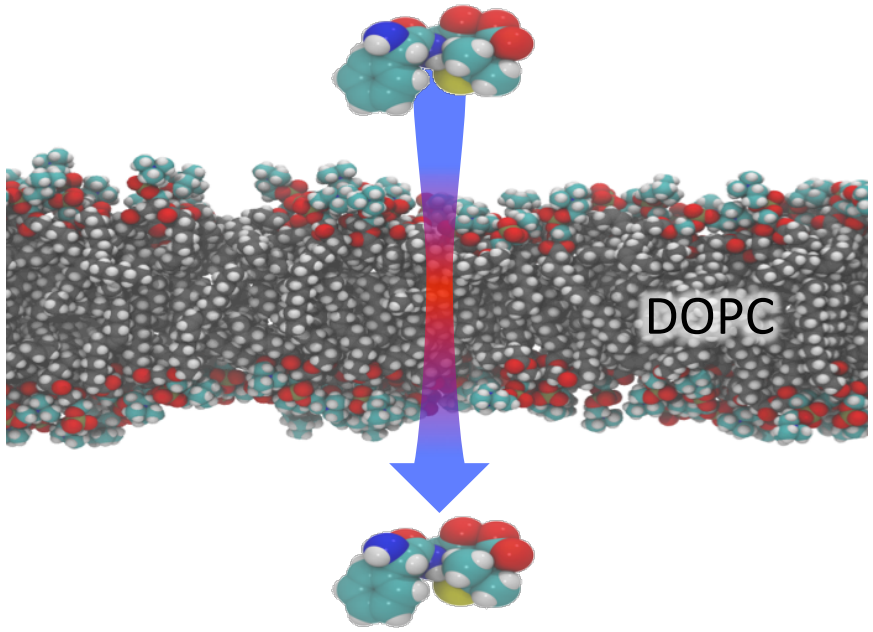
Questions we can address:

- What are the accessible conformations of this system?
- What contacts mediate a small molecule/macromolecular interaction?
- What are likely consequences of small molecule binding?



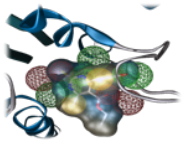
20 ns for ~250,000 atoms
72 hours on 512 Opteron processors

Molecular methods can be used to determine kinetic coefficients

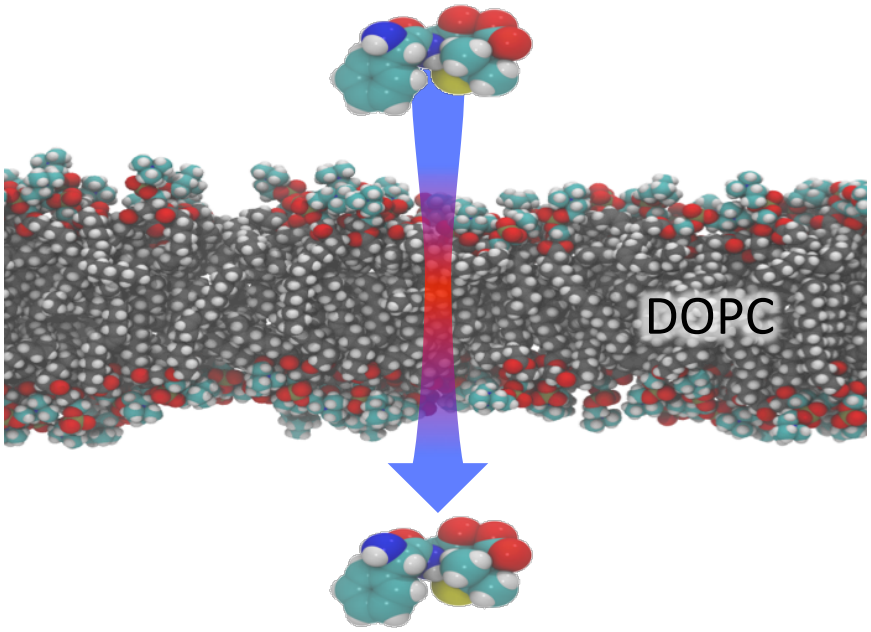


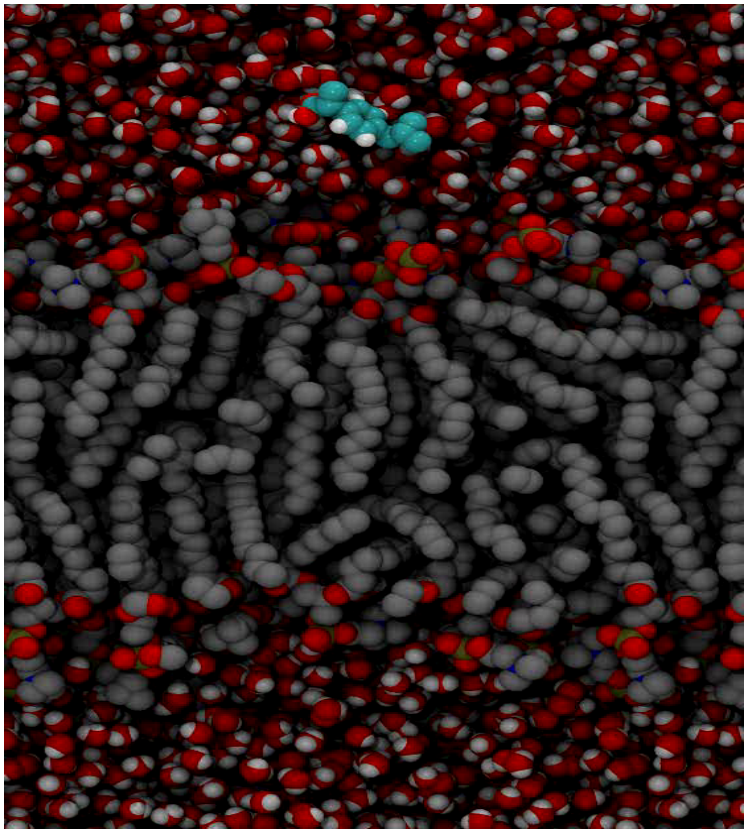
Membrane permeability

- Potential mean force calculations of ligand across a lipid bilayer
- Lipid composition specific for organ
- Validated against experimental results



Molecular methods can be used to determine kinetic coefficients





AMGEN®

Appendix D

Pioneering science delivers vital medicines™

The Future Role of High Performance Computing in Medical Product Decision-making

How will HPC impact the use of non-clinical trials?

Hugo M. Vargas, PhD, DSP
Scientific Director
Integrated Discovery & Safety Pharmacology
Comparative Biology & Safety Sciences

10-Sep-2015

Outline

- **Current Paradigm of Drug Development**
  - Focus: Drug Discovery & Preclinical Development
- **HPC: Opportunities and Challenges**

AMGEN

"Should I kill myself, or have a cup of coffee?"

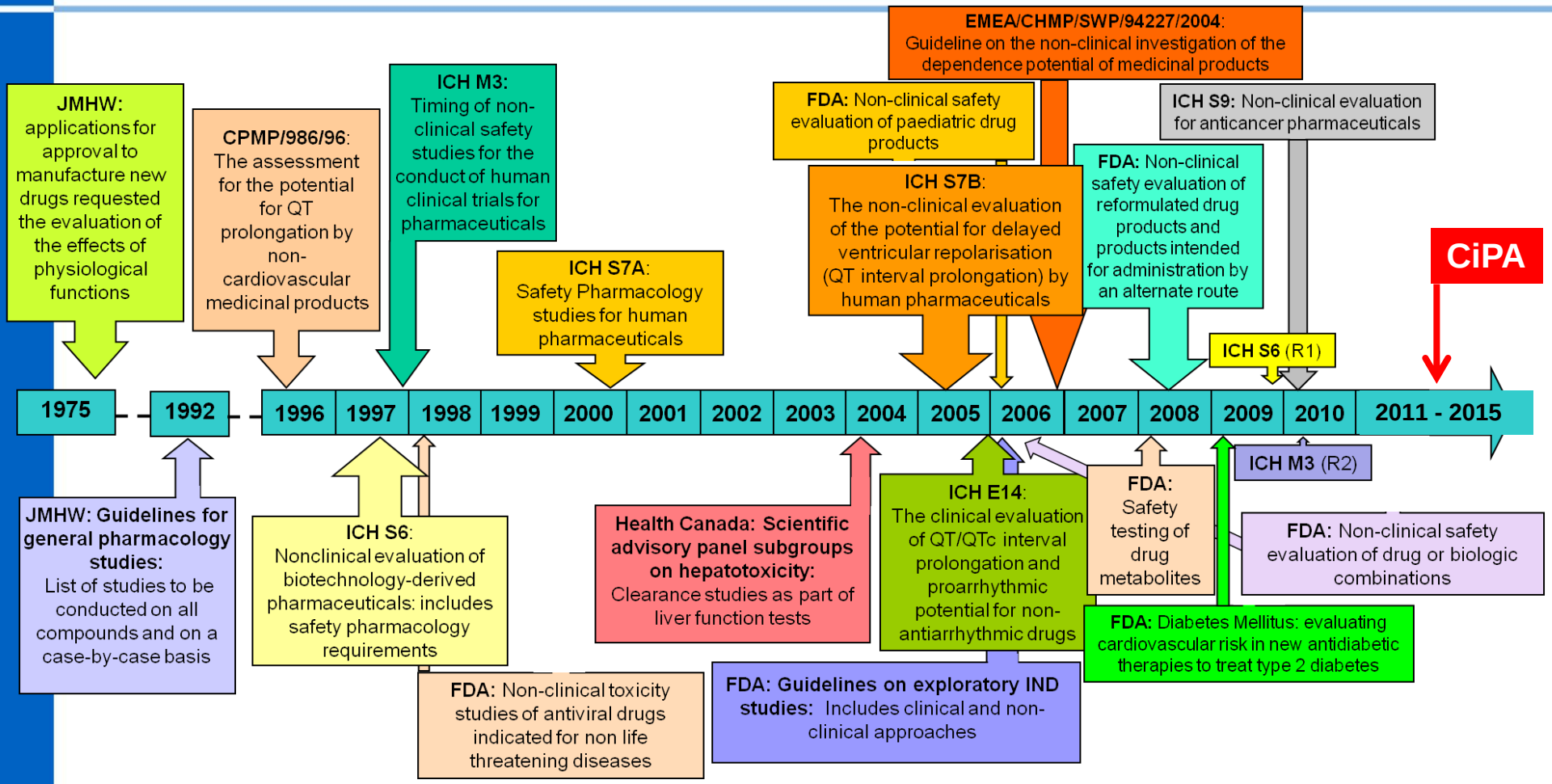
- **Albert Camus** (1913-1960)
  - Author & Playright
  - Nobel Prize (1957, Literature)
    - "…clear-sighted earnestness illuminates the problems of the human conscience in our times"

- **Philosophical Conflict:**
  - "The mundane" vs "The unthinkable"

- **HPC in Drug Development: A Powerful Tool, but How to Implement?**

AMGEN

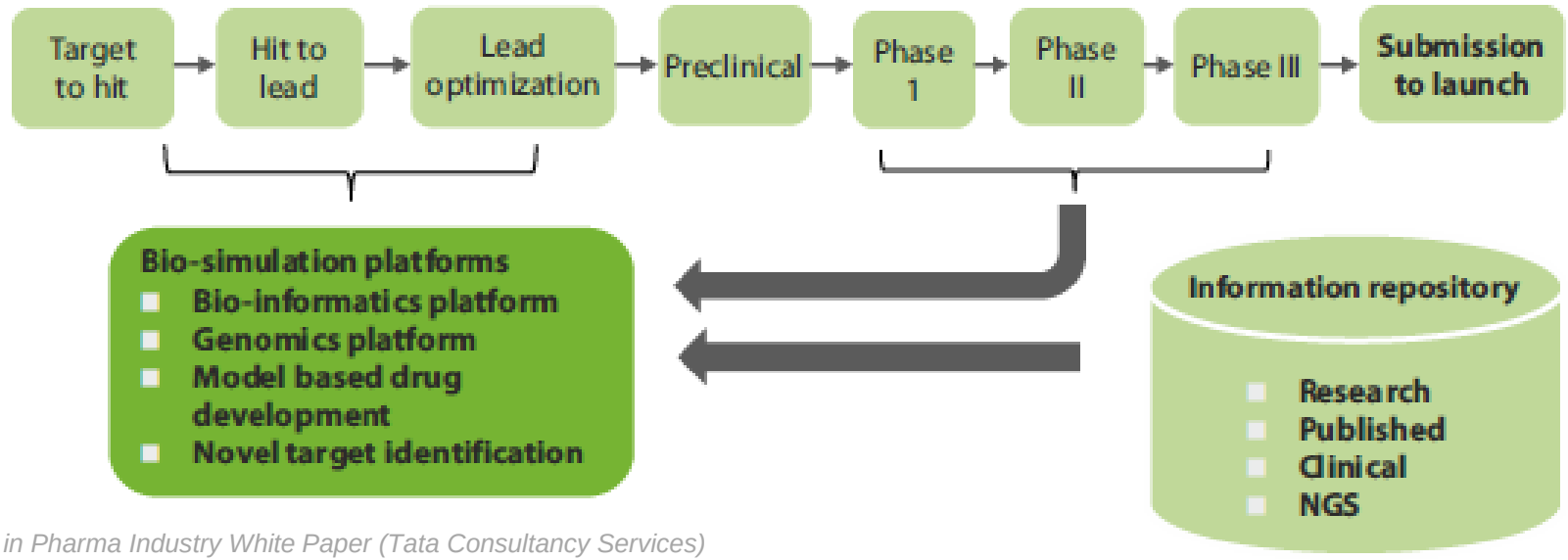
Non-clinical Drug Safety Evaluation (2015): The Regulatory Landscape



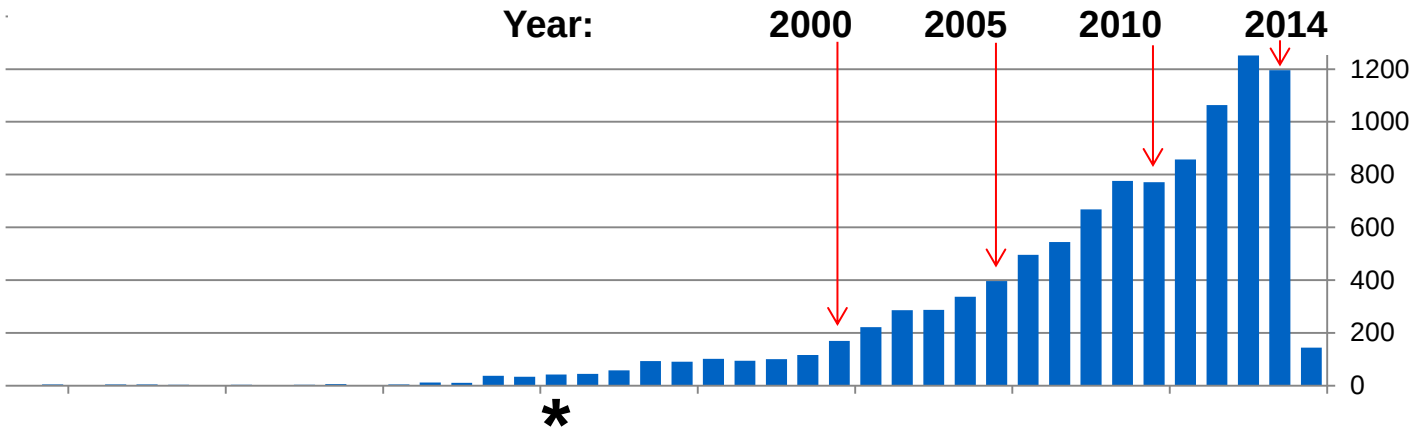
Bass AS, Vargas HM, Valentin JP et al. (2011) *J.Pharmacol. Toxicol. Methods* 64: 7-15

AMGEN

In Silico Opportunities in Drug Discovery:
Nonclinical Phases (small molecule)



HPC in Pharma Industry White Paper (Tata Consultancy Services)

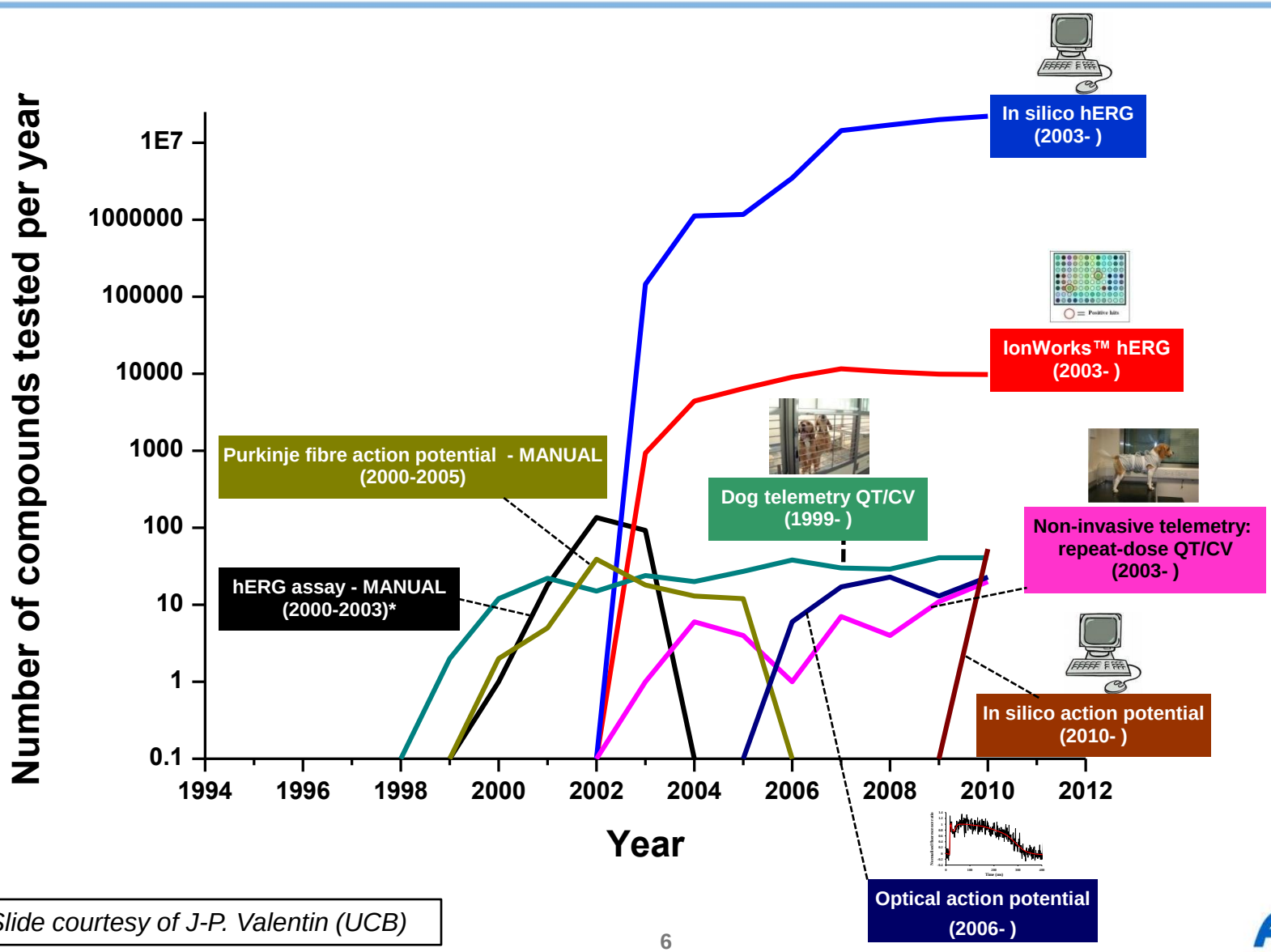


HPC in Drug Discovery
(articles/yr)

Pubmed 1972-2015

HPC: Where to Apply?

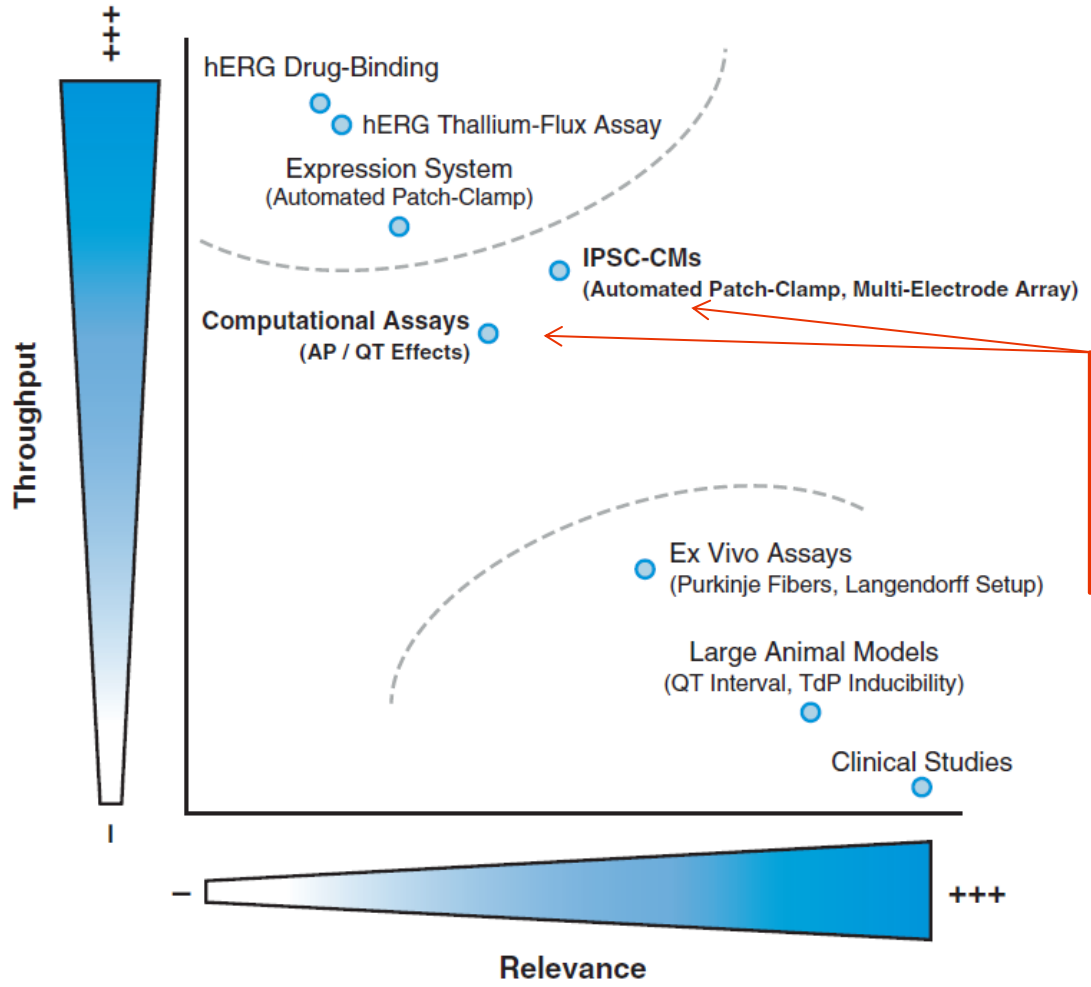
Example: CV/QT Risk Assessment (small molecule)



Slide courtesy of J-P. Valentin (UCB)

AMGEN

Cardiac Safety Assays:
Bridging the Gap Between Models & Clinical Translation



Are these new assays relevant for cardiac safety testing?

Heijman et al. Cardiac Safety Assays. *Curr. Opin. Pharmacol.* 15:16-21 (2014; modified)

AMGEN

Application of SC-Derived CM to CV Safety Assessment: Key Points (SOT-2014)

- **Advantage: Human cardiac tissue surrogate**
  - Potential for better hazard ID, risk assessment & translation
  - Ability to derive risk assessment from susceptible subpopulations: genetic/disease-types

- **Challenge: Scientific validation → Dictate utility & application**
  - SC-CM model: "leading" or "supporting" role
  - Fit for Purpose: needs to be defined

- **Issues: how do SC-CM perform?**
  - Electrophysiology (QTc; QRS) and Contractility effects?
    - Sensitivity, specificity; predictive value
  - Cell type (iPSC vs ES-derived): is one type better?
  - Neonatal vs adult myocyte phenotype: cell maturity factor?
  - Methodology and endpoints: sensitivity?
    - action potential (traditional) vs field potential (multi-electrode array)

Application of ~~SC-[illegible]~~ HPC ~~[illegible]d CM~~ to CV Safety Assessment: Key Points (HPC-2015)

- **Advantage: Human cardiac tissue surrogate**
  - **Potential for better hazard ID, risk assessment & translation**
  - **Ability to derive risk assessment from susceptible subpopulations: genetic/disease-types**

- **Challenge: Scientific validation → Dictate utility & application**

HPC ~~SC-CM~~ **model: "lead" or "supporting" role**

  - **Fit for Purpose: needs to be defined**

HPC

- **Issues: how does ~~SC-CM~~ perform?**
  - **~~Electrophysiology (QTc; QRS) and Contractility effects?~~**
    - **Sensitivity, ~~specificity; predictive value~~**
  - **Cell type (iPSC vs ~~ES~~ derived): ~~is one type~~ better?**
  - **~~Neonatal~~ vs adult myocyte phenotype: cell maturity ~~factor~~?**
  - **Methodology and endpoints: sensitivity?**
    - **action potential (traditional) vs early-after depolarization (EADs), etc, etc.**

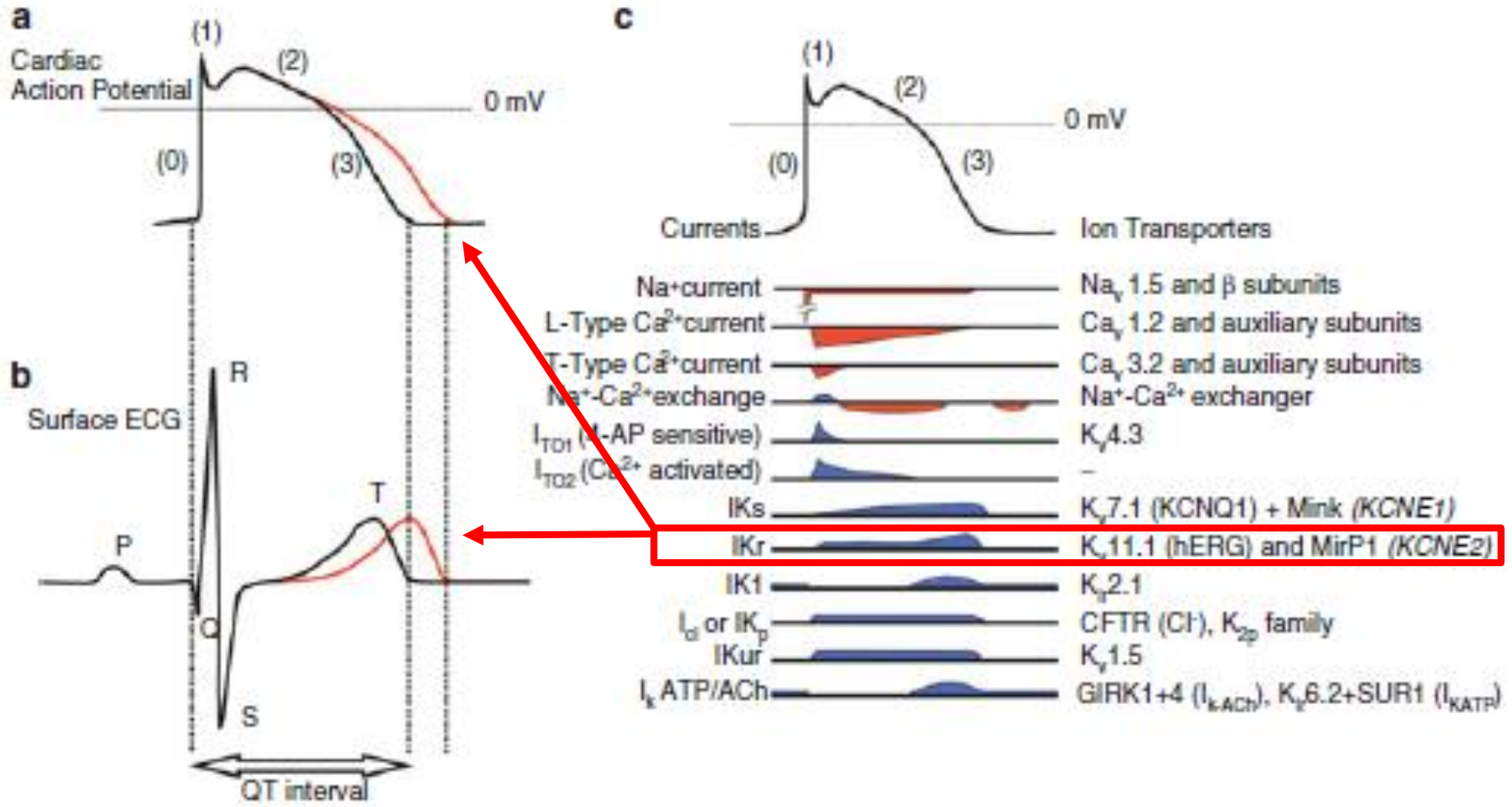
AMGEN

Cardiovascular Safety Pharmacology: A "2020" Vision

An Example of HPC?

AMGEN

Relationship between Cardiac Action Potentials and QT Prolongation



AMGEN

ICH S7B/E14 Guidelines: Intended & Unintended Consequences

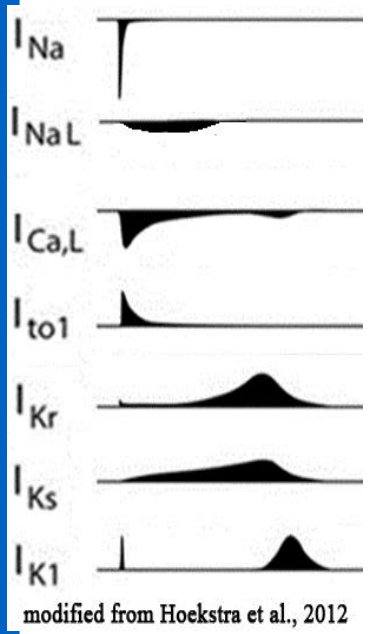
- ***Successful***: no drugs with unrecognized risk being approved or removed from the market

- ***Negative impact*** on drug development
  - Premature discontinuation due to hERG or QT "signal"
    - (Inaccurate) perception of risk leading to drug discontinuation
      - Estimates of up to 60%
  - Development burden: increased costs; labeling
  - Many potentially good compounds never get evaluated in humans due to a hERG effect

AMGEN

Comprehensive *In Vitro* Proarrhythmia Assay: Four Components

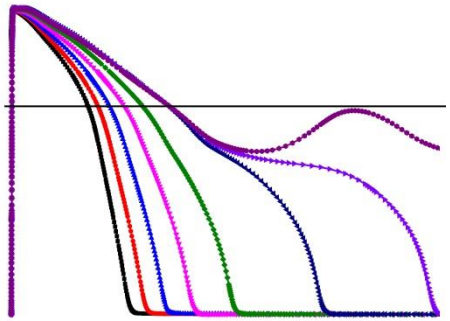
Drug Effects on Multiple Human Cardiac Currents



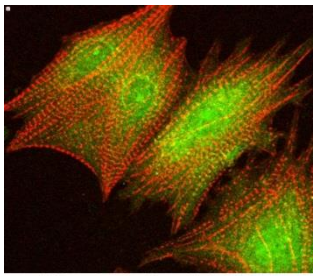
modified from Hoekstra et al., 2012

In Silico Reconstruction Human Ventricular Cellular Electrophysiology

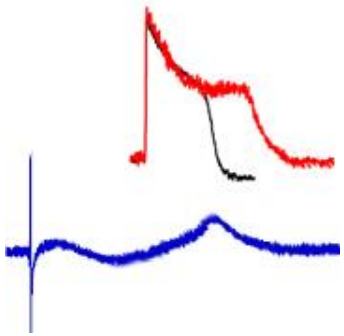
$$I_{stim} = C\frac{dV_m}{dt} + I_m$$



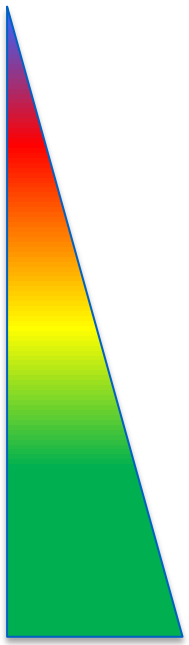
In Vitro Effects Human Stem-Cell Derived Ventricular Myocytes



McEwen Cntr for Regen Med., Toronto



Evaluation of Clinical Drugs for Proarrhythmic TdP Liability



High Risk

Intermediate Risk

Low Risk

AMGEN

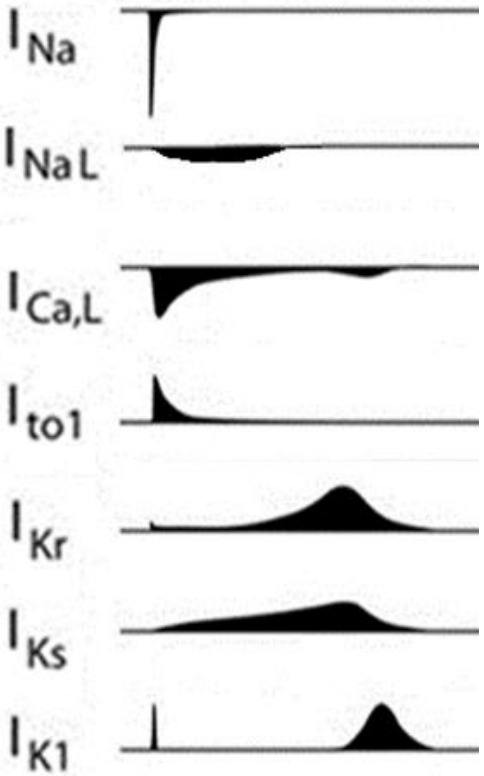
Core Component I: Voltage Clamp Studies, Human Currents, Heterologous Expression Systems

Ion Channel Working Group (SPS):

- Develop seven protocols; standardize testing
  - Outward currents: I_{Kr} (hERG); I_{Ks} (KvLQT1/KCNE1), I_{to} (Kv4.3)
  - Inward currents: I_{K1} (Kir2.1); $I_{Ca\text{-}L}$ (Cav1.2), I_{Na} (NaV1.5; peak & late)

- Establish best practices, reduce bias and variability, enable comparisons of automated platforms across laboratories

- Information on kinetics-, voltage-, and use-dependence to parameterize models (hERG essential)

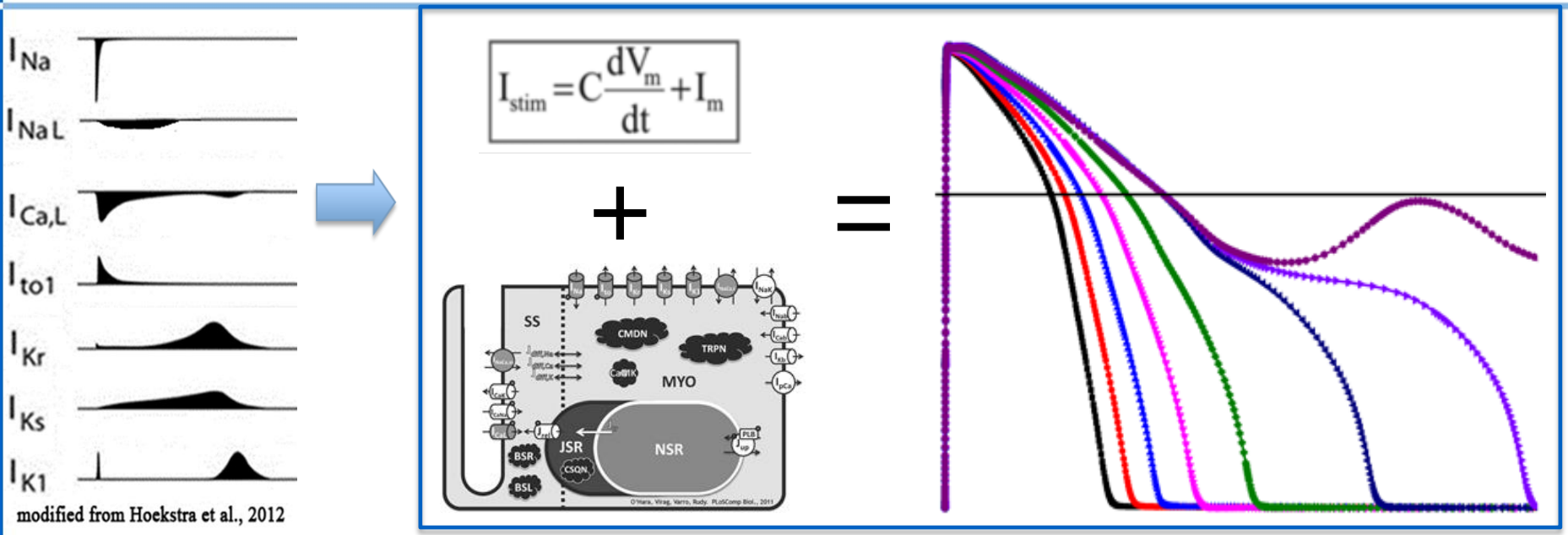


modified from Hoekstra et al., 2012

Robust characterization of drug effects on human currents enables *in silico* reconstructions of integrated responses

AMGEN

Core Component II: Computer Reconstructions of Drug Effects on Human Cellular Electrophysiology



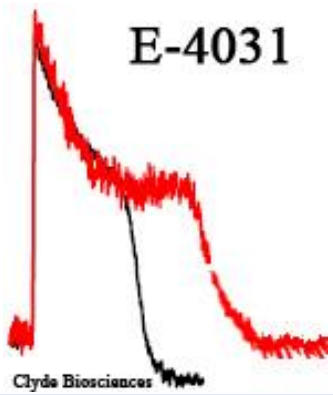
In Silico Group (FDA):

- Multiple currents integrated to describe cellular electrophysiologic effects
- Ability to elicit changes in repolarization instability, early after-depolarizations, reduced upstroke velocity using select model (modified O'Hara-Rudy model)
- Rank integrated responses; compare with clinical examples of TdP risk (low to high risk)

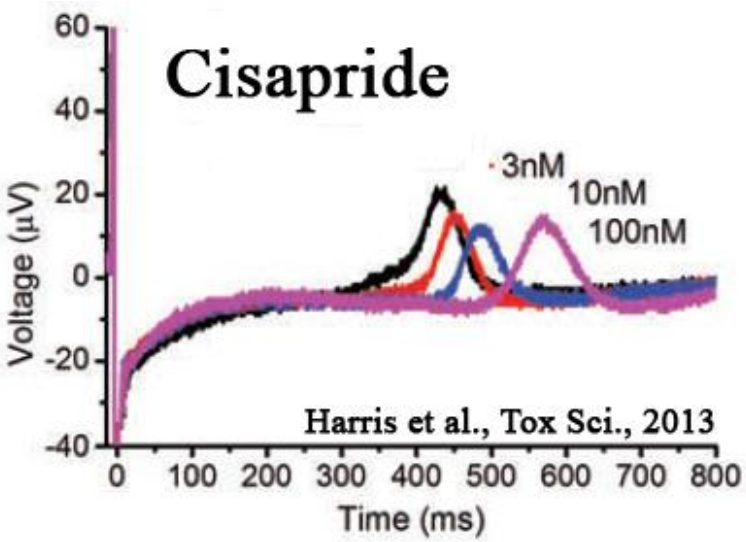
AMGEN

Core Component III: *In vitro* Effects, Human Stem Cell-Derived Cardiomyocytes

Voltage-Sensitive Dye



Micro-Electrode Array



Myocyte Group (HESI):

Verification of *in silico* reconstructions with well characterized human stem-cell cardiomyocytes

13 Site Pilot Study Ongoing

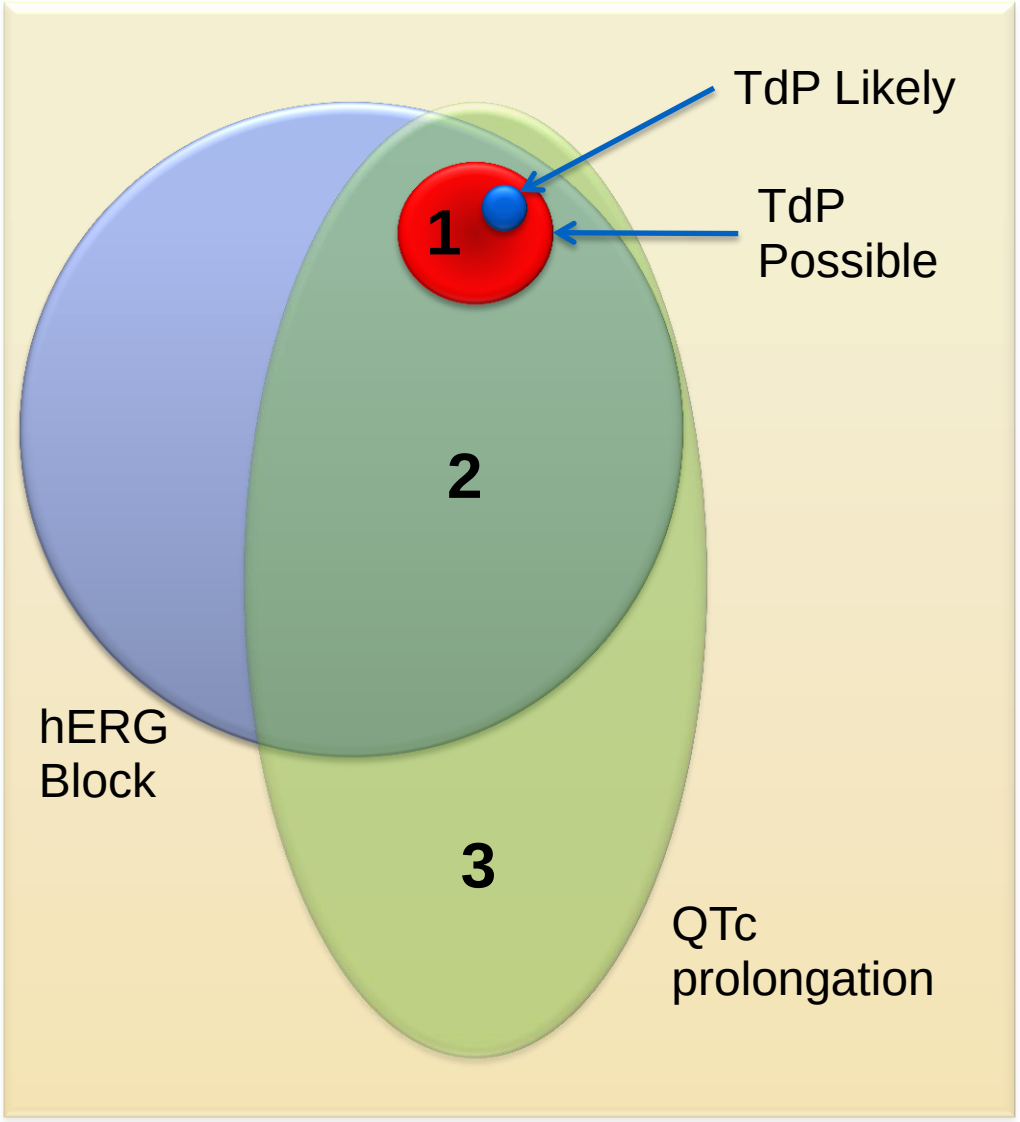
- Microelectrode array (MEA, field potential duration, 4 platforms)
- Voltage-sensing optical (VSO, 4 platforms)
- 3 myocyte types

8 Blinded Compounds

- 4 to calibrate sensitivity (IKr, IKs, INa, ICa)
- 4 as pilot test set

AMGEN

Identifying Different Phenotypes:
Predicting Ventricular Arrhythmia



CIPA Assays Must Differentiate:

- hERG blockers with QTc Prolongation & associated with TdP (**1**)
- hERG blockers with QTc Prolongation *BUT NOT* associated with TdP (**2**)
- Drugs with no-direct ion channel effects with modest QTc Prolongation (**3**)
- Drugs with low pro-arrhythmic risk (3; 4-not shown)

Slide courtesy of D. Leishman (Lilly)

AMGEN

Take Home Points

- **HPC: New Tool with Great Potential**
  - How to apply to drug development?
  - Pro-arrhythmia Risk Assessment (CiPA): Test Case?

- **Is the Juice Worth the Squeeze?**
  - Validation of Models: What is known? What is needed?
  - Resources needed: In silico/HPC/modelling
    - Plus: SME (subject matter expert); multi-lingual
  - How to establish confidence in the output?
    - For business decisions and for regulatory decisions
    - “fit for purpose"
  - Will HPC-derived data supplant/replace current assays used for nonclinical (and clinical) safety assessment?
    - E.g., isAPD (human APD model); Cardioid Project (human ECG model)

Thanks for your Attention!

GIVE ME THE

COFFEE

AND NOBODY GETS HURT

CoffeeNate.com

AMGEN